SHADOWS OF DOUBT:
ON THE PSYCHOLOGICAL FOUNDATIONS OF THE SKEPTICAL PROBLEM

by

John Waterman

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

February, 2015

Abstract: Why is it that we can know so much about the world, but skeptical arguments seem so hard to resist? My dissertation offers a psychological diagnosis of the attraction of skeptical arguments. Just as vision scientists investigate illusion to better understand how perception is successful, I think that by investigating the causes of doubt we can better understand the nature of knowledge.

Most contemporary accounts of the skeptical problem describe it as appealing to our ordinary intuitions about knowledge, and trace its origin to linguistic confusions. I argue that this diagnosis is mistaken on both counts.

The received diagnosis assumes that skepticism is a byproduct of our ordinary epistemic practices. However, this diagnosis is overly reliant on armchair appeals to intuitions about cases by philosophers - individuals who are not ordinary epistemic agents by any means. To better understand the nature of skeptical arguments and their relation to our ordinary epistemic practices, I use the methods of the psychological sciences to empirically investigate how ordinary individuals reason about skeptical arguments, and the factors affect their evaluations. Using this empirical work as a foundation, I argue that the linguistic diagnosis of the skeptical paradox is inadequate because it does not do justice to skepticism's intransigence: linguistic mistakes are easy to recognize, but skeptical doubts are hard to dislodge.

I argue, instead, that the skeptical problem is a byproduct of a conflict between two separate components of our epistemic psychology. Specifically, I defend the idea that we deploy two different heuristic standards of evaluation towards potential beliefs. Toward favored sources, like perception and the testimony of friends, we apply the heuristic *can I believe P?*, and search for evidence that is consistent with P. Toward disfavored sources, like the arguments of rivals, we ask instead *must I believe P?*, and search for possibilities in which P might be false. These two standards, the first fallibilist and the second infallibilist, are inconsistent, and thus lead to the skeptical paradox.

Dissertation Committee: Steven Gross (advisor), Michael Williams, Justin Halberda, Richard Bett, Howard Egeth

Acknowledgements

No project like this is completed without a great deal of help, and I have many thanks to give.

First, I owe an enormous debt of gratitude to Steven Gross for his support and guidance. This essay and my development as a philosopher have only been possible because of his generous insight. I also owe a special thanks to Justin Halberda for teaching me the ropes as I began to experiment with experimental philosophy. My thanks, too, to Michael Williams, whose work on skepticism inspired this project in many ways. I have also been fortunate to have met warm and fiercely intelligent friends in my time at Hopkins. Thank you to Brian Miller, Nick Tebben, Marianna Bergamaschi, and Karen Yan for all your insights and all your time. I owe a special thanks to the members of Language Lab, Derek Leben, Jeff Maynes, and Jon Hricko. You suffered through many of my half-baked ideas and still had the patience to help me improve them. Jon Hricko deserves a special thanks; I would not have started this project without his kick in the pants.

I would also like to thank the administrative staff of the philosophy department for all their hard work and support. Without Alicia Burley, Leslie Bean, and Veronica Feldkircher-Reed nothing gets done, and this dissertation most certainly wouldn't have.

Many parts of this dissertation were delivered at various conferences over the last three years, and I have benefitted enormously from the insights of those audiences and from my commentators. Parts of Chapter 2 were presented at the 2012 meeting of the Southern Society for the Philosophy of Psychology, and I owe the audience there and my commentator David Rose thanks for all their input. Parts of Chapter 3 and 4 were presented at the 2013 meeting of the of the Pacific APA, the 2013 Northwestern Philosophy Conference, the 2013 meeting of the Society for Philosophy of Psychology, the 2013 Meeting of the Buffalo Experimental Philosophy Conference, the 2013 Mid-South Philosophy Conference, the 2014 Southern Society for Philosophy and Psychology, the 2014 meeting of the Central APA, and the 2014 meeting of the Society for Philosophy and Psychology. Many thanks to the audiences at those events and to Wesley Buckwalter, Joseph Schieber, Jeffrey Roland, Dan Korman, Edouard Machery, and Jonathan Weinberg for their incisive comments.

I owe a special thanks to Shaun Nichols and Ron Mallon for offering me an NEH fellowship at the University of Arizona to work on experimental philosophy, and for their generous support since that time. During my time in Arizona I made many friends who contributed to this project. Many thanks to Joshua Knobe, who always checks in to see how things are going, and who always has a helpful idea to offer. My thanks also to Jennifer Nagel, both for answering many questions about the psychology of knowledge attribution, and for offering an approach to philosophy I have emulated. Through my connection with Arizona I also met Joshua Alexander and Chad Gonnerman. We have now collaborated on a number of projects, and they deserve special credit for improving this project.

I owe an enormous debt to my family. I have had the incredibly rare privilege of making my work something that I love. Very few people get that chance. I would never have had it without the love, support, and sacrifice of my parents and my sister. My family became larger in many ways during my time in graduate school, and so I have many thanks to give to Liz Bennett, Joanna Van Orden, Robert Knight, Margaret Chalmers, David Chalmers, John Van Orden, Jenny Van Orden and Jeff Lerman. Thank you all for everything over the years.

An essay like this is made up of words but it is only made with time.

Thank you to Ida, Hazel, and Noel for giving me a reason to finish.

To Barb, who did the most to give me that time, as always, my love.

Contents

Chapter 1

The Intuitive Foundations of the Skeptical Problem

1. Introduction

I'm sure we've all had the experience of going into a department store, finding a shirt that's just the right color, only to get home and see it's a different color than we thought. This happens because colors on the warm end of the spectrum, like red and orange, look quite different under the cool fluorescent lights stores use for reasons of economy. But, knowing this, when we go to the shops it's easy to wonder whether we really know what color things are. What we thought we knew turns out to be wrong just because very often there is no marker or signal of having made a perceptual mistake. Next time you're in the shops, ask yourself: do I know what color this is, or are the lights tricking me?

Another way we make mistakes is through bias. I think by in large, we all try to be objective and treat others fairly. But, there is some straightforward evidence from resume studies that demonstrates how hard this can be. In these experiments researchers ask people to evaluate the credentials of hypothetical candidates, and the only thing they vary is the name of the person at the top. Typically, resumes with male names are rated better than those with female names. Resumes with stereotypically white names are rated better than those with stereotypically African American names. Resumes with stereotypically European names are rated better than those with stereotypically Middle Eastern names. I expect that there is no malice here; people are trying to do their rational, unbiased best. But despite this, their judgment about who is the better candidate is not only a function of credentials, but of group membership.[1] We have tools that can help eliminate bias, and using explicit evaluation procedures and blind review are just a few, but knowing that most people are prone to systematic biases, it's easy to start questioning whether our evaluations are accurate. Next time you're grading a paper, or interviewing a candidate, ask yourself: do I know I'm being fair, or am I being biased?

Yet another case involves autobiographical memory and dreaming. Autobiographical memory is typically very reliable, but it can go wrong on occasion. Jean Piaget describes vivid memories of a moment when his nanny fought off a kidnapper. He only came to revisit those beliefs many years later when the nanny contacted him, and confessed to fabricating the whole episode. Likewise, dreaming can be vivid and realistic, but we're often unlikely to confuse it with our waking life. Nevertheless, as Elizabeth Loftus

[1] R. Steinpreis, K. Anders, and D. Ritzke, 1999; M. Bertrand and S. Mullainathan, 2004; and D. Rooth, 2007.

has shown through a number of studies, it's actually quite easy to prompt people in ways that causes them to believe that elements from dream episodes are in fact memories.[2] Of course we have lots of ways of checking our memory, we can ask our parents and friends, or look at photo albums and Facebook, but once we know how easy it is to acquire false memories, it's easy to wonder whether any specific childhood memory is genuine. Pick a childhood memory, perhaps one when you're alone, and ask yourself: do I know this really happened, or was it part of a dream?

These are uncommon cases for sure, but a feature they share is that the source of the error can be quite difficult to check. Sometimes, we simply can never check. You think you're reading these words on the page in front of you. But how do you really know that what you're seeing isn't the stimulus in an elaborate experiment? How could you tell the difference between *really* reading these words on the page from being a disembodied brain in a vat, who is being *stimulated* to have experiences as of the real world? The case is crazy. It's unreal. But ruling it out seems impossible. There's no marker, no signal we can look for to distinguish reality from a mere stimulation. Without that marker, can we know it's not happening? And if we cannot know that it is not happening, how can we know that what is before us is real?

This is the problem of radical skepticism.

The worries I've just presented, about perception, reasoning, memory, and our knowledge of the external world are all ways of raising doubts through what some call the argument from ignorance (AI). Schematically, AI involves pairing two incompatible hypotheses that would both lead to the same belief. Applied to the perceptual case it looks something like this: I'm looking at a shirt, it looks red, and this is a very good reason to believe that it's red, but…

1. I don't know that the lights aren't changing the color of the shirt I'm looking at.
2. If I don't know they aren't changing the color of the shirt, then I don't know that the shirt is red.
3. Therefore, I don't know that the shirt is red.

[2] G. Mazzoni & E. F. Loftus, 1996.

The argument that I don't know anything about the external world takes the same form. The belief that you have hands is as well attested as any belief that you have: reason, memory, and perception all support the belief, but…

1. I don't know I'm not a brain-in-a-vat
2. If I don't know I'm not a brain-in-a-vat, then I don't know I have hands
3. Therefore, I don't know I have hands.

Of course, my belief that I have hands stands in for just about any belief about the external world, and if it falls, so do the rest. This argument, by easy steps, leads to a conclusion that seems utterly absurd. But it's hard to see where we've gone wrong. If the argument works in the store case, what's wrong with the global skeptic's case?

Hume argued that something in our ordinary epistemic practice invites skepticism. Many contemporary responses to Cartesian skepticism start from the same basic assumption, and attempt to explain what's gone wrong by trying to diagnose why an argument that has such an intuitive starting point can have such an unwelcome conclusion.[3] The tradition sees skepticism as a byproduct of our ordinary epistemic practices, and it this basis in our ordinary thought and talk that is thought to explain the intuitive appeal of skeptical arguments. This approach gives rise to the idea that any response to the skeptical problem must begin with a diagnosis of the intuitions that give rise to it. The approach has been influential, and it's worth seeing how widely shared this characterization of the skeptical problem is.

Here's Barry Stroud,

> When we first encounter skeptical reasoning…we find it immediately gripping. It appeals to something deep in our nature and seems to raise a real problem about the human condition. It is natural to feel that either we must accept the literal truth of the conclusion that we can know nothing about the world around us, or else we must somehow show that it is not true.[4]

[3] Michael Williams (1995) sees this as a distinctive feature of what he calls "the new skeptics."
[4] B. Stroud, 1984: 39.

Here's Thomas Nagel,

> [S]keptical problems arise not from a misunderstanding of the meaning of standard knowledge claims, but from their actual content…[5]

Here's David Lewis,

> I started with a puzzle: how can it be, when his conclusion is so silly, that the sceptic's argument is so irresistible?[6]

Here's Keith DeRose,

> In seeking a solution to this puzzle, we should seek an explanation of how we fell into this skeptical trap in the first place, and not settle for making a simple choice among three distasteful ways out of the trap. We must explain how two premises that together yield a conclusion we find so incredible can themselves seem so plausible to us.[7]

And again,

> [T]he skeptical argument really is powerful... The argument is clearly valid... and each of its premises, considered on its own, enjoys a good deal of intuitive support.[8]

Here's Colin McGinn,

> Philosophical scepticism thus seems endemic to the use of epistemic concepts: to reflect on the concept of knowledge is immediately to question its application…There is something primitive and inevitable about sceptical doubt. It runs deep in human thought.[9]

Here's Bernard Williams,

> [Responding to skepticism] should not be seen as a purely gratuitous demand, a merely obsessional concern with an artificial scepticism. I also think that it is a mistake to see it just as an extra aspiration for knowledge, a superogatory ideal for it which the Western world, at least, has set before itself. We should rather suspect that there is something the notion of knowledge itself which invites this response.[10]

Here's Stewart Cohen laying out the skeptical problem,

> We begin with a paradox – a set of inconsistent propositions each of which has considerable intuitive appeal: (1) I know I have a hand, (2) I do not

[5] T. Nagel, 1989: 69.
[6] D. Lewis, 1996: 561.
[7] K. DeRose, 1995: 3.
[8] K. DeRose, 1999: 2-3.
[9] C. McGinn, 1993: 107-8.
[10] B. Williams, 1978

> know I am not a brain-in-a-vat, (3) If I know I have a hand, then I know I am not a brain-in-a-vat.[11]

Here's Michael Williams,

> A good response to scepticism should be diagnostic and not merely dialectical. We do not want merely to be shown that sceptical arguments go awry: we want an explanation of how they go wrong that also accounts for why they can seem so compelling…scepticism is a problem because the sceptic presents us with apparently intuitive arguments for wholly unacceptable conclusions. In describing sceptical arguments as intuitive, I mean that they seem not to depend on elaborate or contentious theoretical ideas about knowledge or justification. If sceptical arguments were obviously non-intuitive, we could dismiss them (as artifacts of ideas we are not compelled to accept).[12]

Here's Jennifer Nagel,

> If the skeptic is right that there are ways of eliciting negative intuitions about ordinary mundane claims to knowledge - for example, by reflecting on problematic possibilities concerning brains in vats - then the epistemologist who trusts intuitions will have to come up with an explanation of those negative feelings, and a story about how they fit with our more commonly positive feelings about the same judgments.[13]

Here's Ram Neta,

> Skeptical problems can be perspicuously represented as sets of individually plausible but mutually inconsistent epistemic claims.[14]

So, what exactly are the intuitions to which everyone here refers? The best way to see is by breaking down AI step by step.

FH: I have hands

1. I don't know I'm not a BIV (brain-in-a-vat)
2. If I don't know I'm not a BIV, then I don't know I have hands
3. Therefore, I don't know I have hands

Arguments like AI have a focal hypothesis as their target: in this case the well-supported claim (FH) *I have hands.* The first premise introduces what I call a counter hypothesis: in this case an alternative (1) *I am a handless brain-in-a-vat.* If we cannot rule out a counter

[11] S. Cohen, 2001: 87.
[12] M. Williams, 2001: 3-4.
[13] J. Nagel, 2014: 9 mnscpt.
[14] R. Neta, 2003: 396.

hypothesis, so the diagnosis goes, we cannot know that it is false. The idea is familiar from Fred Dretske's relevant alternatives theory. To know that the table before me is red, I need to rule out the possibility that the lights are fooling me. To know that the animal before me is really a zebra, I need to rule out that it's not a cleverly painted mule. The thought behind relevant alternatives theory, and the error intuition that drives it, is the idea that to know that something is the case, we have to rule out relevant alternatives to it. Now, we don't have to rule out all the possible alternatives all the time - that would be an impossible task, we just have to rule out alternatives when they've been made salient in a particular context. So, we can know that the animal before us in the zoo is a zebra…but once the possibility of an error is introduced, it *intuitively feels* like we must rule out that alternative to genuinely know that it is a zebra.[15] Call this the *error intuition*. In the case of premise (1) of AI, once the skeptic's hypothesis has been made salient, we simply cannot rule it out and this is why (1) seems so intuitive.

Well-chosen counter hypotheses imply the falsity of the focal hypothesis. The appeal of the second premise seems to rely on the intuitive principle that knowledge is closed under known implication: the principle that if I know that P, and I know that P implies Q, then I know that Q. If I know that tigers are mammals, and I know that mammals are warm blooded, then I know that tigers are warm blooded. AI involves negation, so it can be somewhat hard to track the principle, but it's easy to think of the following case: if I don't know that my meeting is not on Monday, then I don't know it is on Tuesday. And this is just because if the meeting *could be* on Monday, I can't know it's on Tuesday. But look now at the skeptical hypothesis. If I actually were a brain in a vat, then I wouldn't have hands. If I don't know I'm not a brain in a vat, then I don't know whether I have hands, or whether I don't. Most philosophers have found closure to be a deeply intuitive principle, and one that is hard to reject.[16] To be sure, the principle of closure is not one that many non-philosophers have thought about explicitly; rather the idea is that we all tend to reason in accordance with it, and this is why premise (2) seems intuitive. The *closure intuition* follows from our general rational competence, once we recognize the relationship

[15] F. Dretske, 1970.

[16] F. Dretske does reject it, though, and locates the problem with the skeptical argument at this point. Although I won't focus on closure here, there is widespread support of it. DeRose (1995) thinks rejecting it is "intuitively bizarre;" Fumerton (1987) thinks having to reject closure is a "devastating objection" to any view, and Bonjour (1987) thinks that rejecting it amounts to a *reductio ad absurdum* of a position. CF also R. Nozick (1981) for a wide ranging discussion of closure.

between different states of affairs. The principle of epistemic closure is a generalization of the application to particular cases. The intuitive grip of premise (1) and (2) explains why the conclusion seems to follow automatically. The *error intuition* and the *closure intuition* are the basis for the appeal of AI. I will sometimes simply refer to the appeal of this argument resting on the *Cartesian intuition.*

The full skeptical problem depends on two further intuitions: the *ordinary intuition* and the *intuition of conflict*. The ordinary intuition needs little comment: perception is as reliable a belief forming process as we possess, and our clear knowledge of the facts of experience is as well justified as any belief we might possess. Radical versions of AI often take perceptual claims as the focal hypothesis. While perception malfunctions at the margins, it is generally reliable and the beliefs we form with it are ones we automatically and without question take ourselves to know. Because we are familiar with the ways and the circumstances under which perception malfunctions, when we know we are in compromised conditions we don't automatically endorse appearances as knowledge of our surroundings.

The intuition of conflict is the immediate conflict we seem to recognize between the ordinary intuition that we have hands, and the conclusion of the skeptical argument that contradicts it. The intuition of conflict generally doesn't generate much critical discussion, for it seems as obvious a rational fact as one can muster that "I know I have hands and I don't know I have hands" is a straightforward contradiction. This is not something we need to reflect on to see: reading it, it immediately looks strange, and if anyone were to assert it, it would immediately sound odd.[17] Together, these four intuitions make up the skeptical problem.

2. Past Attempts at Diagnosis

The leading contemporary approaches to the skeptical problem begin by first trying to diagnose the appeal of these intuitions, and then asking whether we must accept them. Let me take a quick tour of some of these. I should mention, though, that the tour is incomplete. My goal is not an exhaustive survey, but a sketch of recent diagnoses that are both prominent and - to mind at least - that rely on tendentious empirical claims about our psychology or knowledge of language. Pragmatic error theory and linguistic contextualism

[17] The intuition of conflict plays a very important role in linguistic contextualism, and chapter 3 will pay close attention to it.

focus on language, while psychological bias theories focus on our psychology. I'll take them up in turn.

Pragmatic Error Theory

Pragmatic error-theorists focus on the intuitive appeal of the *error intuition* and the *ordinary intuition*. As Allan Hazlett develops the idea, "The intuitive appeal of skepticism is, on my view, explained by appeal to the fact that it would be conversationally inappropriate to say that we know in skeptical contexts."[18] Why is this? Hazlett, like many pragmatic error theorists, is a fallibilist about knowledge, and does not think that we need absolute certainty to know particular facts.[19] But as a fallibilist about knowledge, Hazlett must explain why we have the error intuition. For the error intuition seems to rely on the fact that you do need to rule out all possibilities of error to count as knowing - even far-fetched ones like demon deception - once they're salient.

According to Hazlett the error intuition rests on a mistake. We mistake the infelicity of asserting a proposition in a particular conversational context for it being false. This is because, according to Hazlett, when the possibility of error is introduced by a skeptical hypothesis, it makes certain questions salient - if inexplicit - such as "Can I rule the possibility out?" or "Can I be absolutely certain?" or "Can I prove that I'm not being deceived by a demon?" This is the basis of the error intuition. Pragmatic error theorists like to point out that conversations are governed by pragmatic maxims, which facilitate communication. Grice's maxims of relation (be relevant) and quality (be truthful), according to the view, drive our intuitions here.. According to pragmatic error theorists, asserting the focal hypothesis, "I know I have hands," conversationally implies that you can answer one of the above questions, and thus that you can rule out the counter hypothesis even though you can't. So, we have the intuition that we should not assert the focal hypothesis because it would be irrelevant or misleading to do so, and we mistake the intuition that we should not say them for the intuition that they are false. But, according to the pragmatic error theorist, this doesn't matter for what we *actually* know. Even though I cannot answer those questions, I don't need to, because from the normative standpoint knowledge doesn't require me too.

[18] A. Hazlett: 2009: 591-2.
[19] See also P. Rysiew, 2001.

And this is how they deflect the skeptical challenge. In a nutshell: we don't need to answer the skeptic, but because we become confused about language we feel we need to.

Pragmatic error theory has one great theoretical virtue, and one serious flaw. Its virtue is that it can explain our ambivalence toward the conclusions of skeptical arguments: why we are sometimes gripped by them, while other times we are not. What is pragmatically assertable in a conversational context can shift quite readily, and this explains why the same claim, "I have hands," can feel so intuitive in one context and unintuitive in a skeptical context.

Its flaw is that it must attribute to us two kinds of systematic errors. If knowledge doesn't require (i) ruling out all alternatives, (ii) proving we're not deceived, or (iii) establishing that we can be certain, then why does discussion of what we can know always pragmatically imply these question? If they're not genuinely connected, then it's strange that they seem to so regularly co-occur. But the pragmatic error theorist is committed to (i) - (iii) being reliable conversational implicatures. The second is that it must posit that we consistently mistake what is pragmatically unassertable for falsity. By in large, it is better to solve problems without attributing widespread error to all involved. On final worry, which I'll discuss further in a later chapter, is that pragmatic approaches like this are deeply wedded to the use of conversational dynamics to explain the ambivalence of skeptical doubts. The difficulty with this strategy is that entertaining skeptical doubts is something we can do alone. The kinds of conversational misunderstanding they posit just don't apply to solitary reflection.

Linguistic Contextualism

Contextualism is based on two empirical claims. The first is the linguistic claim that the verb "know" is covertly context-sensitive. The second is the psychological claim that ordinary speakers exhibit a selective "semantic blindness" to the context-sensitivity of "know," and that this blindness explains the appeal of the skeptical problem.

Contextualism comes in many forms. Norman Malcolm suggested "know" is ambiguous between and ordinary sense and a philosophical sense with exceptionally high standards. More recently contextualists like Stewart Cohen have claimed that as the standards of accuracy in a context vary, so too does the required threshold for justification in

that context.[20] Others, like J.L. Austin, David Lewis, and Jonathan Schaffer claim that knowledge is sensitive to the range of contrasts that are relevant in a context.[21] But perhaps the most influential and well-developed view, due to Keith DeRose, is that knowledge depends on the contextually determined range through which one must track the truth.

As DeRose formulates it "know," a verb, has a semantics similar to scalar absolute adjectives like "flat," "empty," and "tall."[22] For example, the sentence (1) "Tom is tall" might be true in the context of the elementary class where he teaches, while "Tom is tall" might be false in the context of the basketball league where he plays after work. Because "tall" is context-sensitive, the semantic contribution it makes to the sentence that embeds it, and so the truth conditions of the whole sentence, can vary with context.

Contextualists claim that the standards for knowledge can shift from one conversational context to another. So, amidst the low epistemic standards of everyday conversation, it may be true that (2) "Tom knows he has hands," but under the higher epistemic standards of a philosophy seminar, a token of the same sentence might be false. Contextualism is, consequently, a concessive solution to the skeptical problem. It claims that we find the error intuition compelling because the skeptic ratchets up the standards in such a way that it is true that we don't know that we're not a brain in vat. Contextualists also endorse the closure intuition, and so, at high standards the skeptic's argument is sound. What the contextualist also claims is that in the lower standards of ordinary contexts, we do know we have hands, and this explains the appeal of the ordinary intuition. Contextualists often motivate their case for the context-sensitivity of knowledge ascriptions by pointing to pairs of cases. Consider the following cases adapted from Steward Cohen.[23]

Normal Case	John A. Doe is in a furniture store. He is looking at a bright red table under normal lighting conditions. He believes the table is red.

Table 1.

[20] S. Cohen, 1988, 1999.

[21] J. L. Austin, 1946; D. Lewis, 1996; J. Schaffer 2005.

[22] See P. Unger 1975; S, Cohen, 1999; DeRose, 2009; see also Lynn Murphy (2010: 233) for a discussion of classical examples of scalar absolute adjectives.

[23] See S. Cohen, 2005; J. Nagel, 2010; J. Alexander, C. Gonnerman, & J. Waterman, 2014. These vignettes will form the foundation of some of the experimental work I pursue in Chapter 4.

Does John know that the table is red? According to contextualists the intuitive, and obvious, answer is yes. But now consider the following case.

Skeptical Case	John B. Doe is in a furniture store. He is looking at a bright red table under normal lighting conditions. He believes the table is red. However, a white table under red lighting conditions would look *exactly* the same to him, and he has not checked whether the lighting is normal, or whether there might be a red spotlight shining on the table.

Table 2.

Now ask yourself the same question. Does John know that the table is red? According to contextualists, the intuitive, and obvious, answer is no. Our intuitions shift, even though John's epistemic situation hasn't changed at all. No new information has been brought to his attention; all that's happened is that an unrealized possibility of error has been made salient. For contextualists, our intuitive assessments of pairs of cases such as these are both the empirical evidence for the claim that "know" is context-sensitive, and an explanation of the appeal of the error intuition behind arguments like AI.

According to contextualists, we feel gripped by the skeptical problem because we have a semantic blindness to the context-sensitivity of "know," which causes us to feel the intuition of conflict. Recall sentence (1) above. We haven't contradicted ourselves by saying "Tom is tall" in one context (where he's a teacher) and saying "Tom is not tall" in another context (where he plays basketball), nor do we intuitively feel like we've contradicted ourselves, because we're sensitive to the context-sensitivity of "tall." But contextualists contend that we are blind to the context-sensitivity of "know" and this is why the claims "I know I have hands" (uttered in an ordinary context) and "I don't know I have hands" (uttered in a philosophy seminar) seem intuitively contradictory.

Contextualism is concessive. It allows the skeptic a limited victory at high standards, so it gives up more ground than pragmatic error theory. Like pragmatic error theory it has the virtue of offering and explanation of the ambivalence of skepticism, but at the cost of attributing widespread error.

Psychological Bias Theories

Psychological bias theories locate the error not in the misunderstanding of our own language, but in the psychology underlying our practices of epistemic assessment. Some invariantists like Timothy Williamson and John Hawthorne argue that the implausibility of skeptical hypothesis plays an essential role in understanding the grip of skeptical arguments. They claim that it is because skeptical hypotheses are psychologically arresting, we overestimate the threat to they pose to ordinary knowledge, and this explains the error intuition. Both point to what's known as the "availability heuristic" in support of this diagnosis. Williamson contends,

> One effect of fictional violence on television is to make viewers overestimate their chances of being the victims of violent crime: they suffer an illusion of danger. Might not an illusion of epistemic danger result from exposure to lurid stories about brains in vats, evil demons, painted mules, or gamblers who bet the farm?[24]

John Hawthorne concurs, offering a very similar diagnosis of the origin of salience effects,

> When certain non-knowledge-destroying counter possibilities are made salient, we overestimate their real danger; as a result, we may find ourselves inclined to deny knowledge to others in cases where there is in fact no real danger of error.[25]

According to Daniel Kahneman and Amos Tversky we tend to use the ease with which a possibility is imagined as a proxy for determining the likelihood of the event.[26] The classic example of the availability heuristic at work comes from assessments of the danger posed by different forms of travel: people tend to overestimate the danger of flying and underestimate the danger of driving because plane crashes often occasion breathless media attention and graphic visuals, and this makes them much easier to remember. Williamson and Hawthorne agree that premise (1) of AI is intuitive, and their account has the virtue of safe guarding our knowledge of everyday facts. As with pragmatic error theory and linguistic contextualism, their diagnosis is premised on empirical claims: the first is that we find radically skeptical

[24] T. Williamson, 2005: 226.
[25] J. Hawthorne, 2004: 164.
[26] A. Tversky & D. Kahneman, 1973.

versions of AI intuitive, and the second is about the psychological process that explains why we find it intuitive.

Jennifer Nagel offers an account that is similar in spirit. Nagel's primary focus is explaining the variation in our intuitive willingness to attribute knowledge, especially in the kinds of cases we saw contextualists appeal to above (Tables 1 & 2), not diagnosing the skeptical problem as a whole.[27] She agrees that raising a possibility of error intuitively undermines our willingness to ascribe knowledge. But, she thinks this intuition results from a cognitive bias known as epistemic egocentrism, whereby people have difficulty distinguishing their representation of affairs from the representations of others. Nagel argues that because of epistemic egocentrism, when a possibility of error is introduced about an agent's belief, we treat our concerns as theirs, and we penalize them for failing to respond to those concerns appropriately. Taking the John Doe cases as an example, although the agent in the vignette isn't aware of the error possibility - it's been aired by the omniscient narrator, Nagel proposes that we project awareness onto the agent, and then penalize him for forming a hasty belief.[28] Nagel's proposal is attractive to invariantists because it offers a way to explain salience effects as the *product* of a bias. But just as with the theories we've looked at above, it is a diagnosis that is based on two empirical claims: one about the intuitive basis of the skeptical problem, and another about the psychological basis of that intuition. A difficulty facing Nagel's theory, as we'll see, is that it only works for knowledge ascriptions to others, i.e. the practicing of assessing the mental states of others. We penalize people for failing to address a worry that we (wrongly) think they should address. But this kind of model does not do a good job of explaining the first person appeal of raising an error possibility. When an error possibility is made salient to us, about something we believe we know, there is no question of projection, but it can still seem to undermine our commitment even when we're not being hasty.

3. Empirically Informed Diagnosis

While the dominate approaches to understanding the skeptical problem accept that it is an intuitive one, by no means everyone agrees, or believes that the problem lies in our ordinary epistemic practice. William Lycan, for one, denies this concessive starting point,

[27] Though I believe that this is a project that she is laying the groundwork for.
[28] J. Nagel, 2010: 303.

> [I]t is often claimed that the skeptic's premise is not made-up philosophy stuff, but is already – if not analytic, unavoidably involved in assessment of our knowledge of the world, or itself common sense – tacitly accepted **by philosophers and by ordinary people everywhere**. On this view, the skeptic is not dragging in some disputable philosophical premise, but is launching a fair and sound ad hominem against the nonskeptic. Wrong, and/or question-begging, again: Who says we all accept the premise? The historical fact is that we do not all accept it, and we are not all being obviously irrational. At the very least, argument is needed, and the argument itself is sure to have some distinctively philosophical premise.[29]

So, just how intuitive is the skeptical problem? Does it really rely on nothing more than our ordinary concept of knowledge and our ordinary epistemic practices? I think there is very good reason to think that it is far less intuitive, and far less plausible a threat to ordinary knowledge than is widely thought. Many people disregard radical skepticism when they first encounter it. As philosophers, we think our task is to motivate it. But I think this rejection is a feature of the skeptical problem, not a fault in ordinary understanding. Initially rejecting skeptical arguments is different from the ambivalence of skepticism that I described earlier. To those of us who think the skeptical problem is a problem, the threat comes and goes, as Hume observed depending on whether we are reflecting on the problem, or just going about the business of everyday life. And although they may be related, this is quite different from resisting the problem from the outset, or never finding it threatening. I believe there is more structure to the skeptical problem than previously thought. I believe there are more individual differences in response to the skeptical problem than previously thought. A genuine diagnosis of the skeptical problem is only going to be possible if this variation is accurately described. The goal of this essay is to understand the problem of radical skepticism and the instability of its appeal by developing an empirical description of the problem that can serve as a firm foundation for its diagnosis.

Consider how most of us respond to conspiracy theories: very few people really think that George W. Bush orchestrated the 9/11 attacks, or that NASA faked the moon landings. But these are logically possible counter hypotheses about how the world is, and they're views with real advocates. The challenge conspiracy theories pose to our everyday knowledge has same structure as radical skeptical arguments. The point of radical skepticism

[29] W. Lycan, 2001: 46-7. Emphasis added.

is not to convince you that you're a brain-in-a-vat, but to show that the inability to rule out the hypothesis seems to undermine ordinary knowledge claims. But why don't we feel the same pull with conspiracy theories? I know Neil Armstrong landed on the moon, and absurd conspiracies are no threat to my knowledge.

And when someone says they don't find the skeptic's argument intuitive, what should we say? Stewart Cohen has a constructive stance. He also sees the skeptical problem as an intuitive one, and recognizes its ambivalence: "A robust feature of those intuitions is that we tend to vacillate between worrying about whether we know anything, worrying that skepticism is true, and thinking that skepticism is just crazy."[30] The constructive task of any diagnosis and response to the skeptical problem is to explain, or "explain away the skeptic within ourselves."[31] However, the intuitive characterization of AI has proceeded from the armchair, and there is a real reason to worry that it is an incomplete description as a result.

Recently, experimental philosophy (sometimes known as just x-phi) has emerged as a tool for studying intuitions about all manner of philosophical questions in a more rigorous way than through introspection. Some of the early results have been striking. For instance, there is perhaps no more influential thought experiment in epistemology than the Gettier case. The consensus intuition among philosophers is that the agent in the Gettier case doesn't know, and this commitment is as widely held as any in philosophy. But, strikingly, when you turn to the question over to ordinary people - the folk - it turns out that they *do not* tend to share the philosopher's intuition. The results are complicated, and the resulting debate is something I'll touch on throughout the course of the essay, but the worry is clear: ordinary intuitions may be quite different than philosophical intuitions.

Experimental philosophy, then, has an important lesson for the study of skepticism: different groups may have different intuitions. If the problem of skepticism is one that emerges out of our ordinary epistemic practices, and not from a specialized philosophical conception of knowledge, then the lesson is clear. Diagnosis cannot begin by describing the skeptical problem on the basis of the intuitions of a small, highly idiosyncratic group. Philosophers' intuitions are unlikely to be representative of ordinary epistemic practice. Moreover, any diagnosis that does not explain the ambiguity of the skeptical problem - when and why it grips us, and when and why it does not - is deeply incomplete. The transience of

[30] S. Cohen, 2001: 94-5.
[31] S. Cohen, 1999: 63.

the skeptic's threat is as important as the threat itself. Indeed, it's an open possibility that the skeptical problem is only a philosopher's problem, one that only emerges from their extraordinary epistemic practices.

Michael Williams recognizes the centrality of the threat's transience, and he faults those who are too willing to concede that the appeal of the skeptic's challenge is evidence that it emerges from nothing more than our ordinary epistemic practice, saying,

> To hold that skeptical doubts are natural is to deny that the sceptic depends essentially on distinctive theoretical commitments not clearly implicit in our ordinary handling of epistemic concepts…[But]…skeptical arguments depend essentially on theoretical commitments that are *not* forced on us by our ordinary ways of thinking about knowledge, justification, and truth.[32]

I would drive the wedge deeper: skeptical doubts can seem intuitive, at least sometimes to some people, but even if a compelling intuition is embedded in our ordinary practices, it doesn't follow that it embodies a commitment that we must accept, upon recognizing it for what it is. Some people find conspiracy theories deeply intuitive, but once we recognize the psychological foundations of those intuitions, and the principles they operate on, it is an open and further question whether we would endorse them.

Williams offers Zeno's paradoxes of motion as an example of the point. It's an analogy I'd like to explore, and expand. Zeno's paradoxes of motion are intuitive, but they lead to a conclusion that is deeply implausible. Trying to respond to the paradox, though, quickly leads to substantial questions about infinity and the divisibility of space that are clearly theoretical questions in fundamental physics.[33] The analogy is apt, in my view, because it points to the ways in which certain paradoxes are intricately linked with our evolved psychology. Michael McCloskey's now widely influential work on folk physics illuminates the point.[34] We have certain intuitions about the mechanics of physical objects that are at odds with classical mechanics. The explanation for this is just that our psychology evolved to respond to the kinds of problems our ancestors confronted. Where our intuitions about mechanics depart from actuality is in part a product of those differences having little adaptive significance. Our psychology evolved to solve problems, and those

[32] M. Williams, 1996: 31-2.
[33] M. Williams, 1996: xviii.
[34] M. McCloskey, 1983.

solutions didn't have to track the truth, they only needed to be adaptive. But even if they evolved to solve specific problems, it doesn't mean that the mechanisms that deliver intuitive verdicts don't embody specific theoretical commitments.

There is now a wide literature in psychology that shows our ordinary rational practices and forms of assessment reflect a tacit reliance on principles that deviate from normative standards.[35] Perhaps one of the most well known illustrations is the Wason selection task. If you present people with four cards, each with a letter on one side and a number on the back, and ask them which they need to flip over to test whether a rule is true of the cards, they invariably focus on possibilities that confirm the rule and ignores possibilities that would disconfirm it.[36] We also tend to exhibit *belief bias* when evaluating arguments: in evaluating the strength of an argument, we tend to only focus on the conclusion and not the form. We tend to accept arguments when they have conclusions that we antecedently accept, even if the argument itself is invalid.[37]

These are just two examples of many. There is also a wide literature on assessment bias. We've already seen the availability heuristic, and epistemic egocentrism. Another set of striking findings comes from the field of motivated cognition.[38] We tend to apply different standards of evaluation to claims, depending on their content. When a claim seems to fit with our overall picture of the world, or if it seems to cast us in a positive light, we tend to ask the question "can I believe this?" and we seek justification for the belief. When a claim doesn't fit with our overall picture of the world, or if it seems to cast us in a negative light, we tend to ask the question "must I believe this?" and we seek out reasons to reject the belief. For example, when people are asked to compare their driving ability to the general population, the vast majority place themselves in the top quartile. But of course most people cannot be above average.[39] Looking at the negative side, Dan Kahan has recently done a

[35] The heuristics and biases tradition is a large one, and there are many excellent summaries of the results. Daniel Kahneman recently published a popular account of the field, and his seminal work in it, in *Thinking Fast and Slow*. See Edward Stein, *Without Good Reason*, for a philosophical appraisal of the significance of this tradition.

[36] For example, if the cards were 7, E, 4, and F; and the rule under evaluation was "if there is a vowel on one side, there's an odd number on the back," then people invariably tend to flip "7", but ignore "4".

[37] J. Evans, 2003. But then just compare this folk strategy of evaluation to the consensus evaluation of AI and Moore's argument. Moore runs AI in reverse, claiming that: I know I have hands, and if I know I have hands I also know I'm not a BIV, so I know I'm not a BIV. But the consensus has been that while both arguments are valid, there is something wrong with Moore's - it's not persuasive despite supporting commonsense.

[38] Z. Kunda, 1990; T. Gilovich, 1991.

[39] T. Gilovich, 1991.

great deal of work on group identification and openness to evidence. He's found that even controlling for scientific literacy, people's group orientation (liberal or conservative) tends to predict their openness to evidence about contested topics. For instance, if you're a liberal, you tend to accept evidence that shows climate change is anthropocentric while discounting evidence that it is not.[40] If you're conservative, the opposite is true. The same pattern of assessment emerges when liberals and conservatives consider things like vaccinations and autism, evolution and human origins, and weapons of mass destruction and the second gulf war. Of course liberals and conservatives have different beliefs, but what's striking is that both groups exhibit the same pattern of responses to confirmatory and disconfirmatory evidence. What's more, there is an emerging body of evidence that the longer a person considers evidence for a particular claim (or against it) the harder it becomes for them to even imagine that the view could be false.[41]

The heuristics and bias literature is large, and growing larger. Here I've given but the briefest of tours. What these biases have in common are two important features: they are automatic assessments, and they rely on principles that we would not endorse as normatively appropriate, if they were explicit. I'll eventually return to the idea that the grip of the skeptical problem may depend on cognitive biases, but first I want to point out something quite simple. Discovering these biases, and by "discovery" I mean establishing that that these patterns of reasoning and assessment are real patterns of everyday reasoning, relied on empirical work. Establishing that a pattern is a pattern depends on collecting evidence from a large, diverse population, and patterns of bias like the ones I've discussed cannot be established as a pattern by a priori reflection. We can reflect on our own thinking, and speculate about whether it exhibits a pattern that other people share (or don't), but to really see if it's out there, you've got to do the empirical work.

Just as Zeno's paradox may depend on tacit commitments about divisibility that are part of our folk physics, I believe that the skeptical paradox may depend on the tacit principles of our folk epistemology. Uncovering the structure of those commitments is an empirical project. The reason is simple, and it's the same reason uncovering biases like those above is an empirical project. Normative theories aim at consistency, but nature faces no such constraint. Nature is a satisficer: it only aims at good enough. So understanding our

[40] D. Kahan, 2014.
[41] E. Mandelbaum, 2014.

rational capacities and our epistemic psychology really is an empirical project. But the advances of the heuristics and biases tradition have been aided by the fact that we have well developed normative theories to compare our psychological competencies against. Because of this we're well placed to say what is a bias and what isn't.[42] We're not so well placed when it comes to epistemology.

I believe moral psychology offers apt comparison in this respect. In the last decade moral psychology has made major advances in understanding our moral cognition by undertaking an empirical turn. It does not aim to replace moral theory, but to explain the underlying mechanisms that give rise to our moral intuitions, intuitions that in turn drive a great deal of theorizing. By taking the empirical turn, moral psychology has provided significant evidence that our moral cognition is not only organized in advance of experience, but also that those structures are complicated, intricate, and sometimes inconsistent. A part of our moral cognition is innately specified because living in groups and managing the conflicting aims of their members was one of the key challenges our ancestors faced. It is inconsistent because those challenges changed over time, and as our species faced new challenges it developed new capacities. Our moral brain is a kludge, not an elegant machine.

We should expect nothing less with epistemology. Although most reflection in epistemology on our underlying psychology tends to focus on our "concept of knowledge" or how this "concept" might depend on our folk psychology, no one in moral psychology believes moral cognition depends on a single concept or system. Moral cognition is subserved by a diverse set of systems because it evolved over long epochs to solve many problems. Just as living with others was a major adaptive problem, so too was extracting information from the environment, from perception to social learning, and ultimately to reasoning. There is every reason to think that our epistemic psychology is just as messy, just as kludgey, and that understanding its workings will depend on the same kind of empirical undertaking.

Returning to the skeptical problem, my advocating an empirical stance in diagnosing it does not mean that experiments can take the place of evaluating the skeptic's position. It just means that an adequate diagnosis should proceed from an accurate description of the problem. And there are good reasons to think that the description of the problem I

[42] Here I'll ignore the distinction between normative rationality (truth conducive and preserving procedures) and ecological rationality (procedures that are conducive to survival).

canvassed above is both inaccurate and incomplete. The leading tradition of diagnosis, from Austin and Malcolm, to DeRose and Lewis, has traced the origin of skeptical doubts to linguistic confusions.[43] This is a mistake. Linguistic errors - misunderstanding, misattribution and miscommunication, and all manner of pragmatic mistakes - are all too easily recognized to explain the enduring attraction of Cartesian skepticism. They do not explain its ambivalent grip. They do not explain the individual differences in its attraction. Linguistic explanations misconstrue the problem as a conversational confrontation rather than as a personal encounter. Faced with these limitations, we have a reason to look elsewhere for a better understanding of the skeptical problem.

4. Psychological Foundations

In this essay I'll defend the idea that Cartesian skepticism has a psychological foundation. The appeal of arguments like AI is a product of our epistemic psychology. I'll argue we have two major strategies for assessing claims. One follows an inferential principle that aims at confirmation: it seeks reasons for accepting a claim. A second follows a principle that aims at disconfirmation: it seeks reasons for rejecting a claim. The skeptical problem is ultimately rooted in a conflict between the outputs of these systems. This diagnosis of the skeptical problem has remained elusive because the description of the problem has been overly reliant on *philosophical* intuitions about knowledge and knowledge ascriptions, to the exclusion of *folk* intuitions about the same. Consequently, the diagnostic tradition has been overly concerned with those already in the grip of doubt, and the many attempts at describing the problem within it have ignored the many factors that are important preconditions for pushing doubt to Cartesian extremes. I take seriously the fact that some people are unmoved by skeptical arguments. I take seriously the fact even those who find skeptical arguments intuitive are most often unmoved by them when, as Hume says, "we leave the study." No diagnosis of skepticism can be complete if it ignores those unmoved by skeptical intuitions because it will be unable to accurately describe the isolated conditions that prime those intuitions.

The working hypothesis of my research program is that skeptical arguments are a powerful analytical tool for understanding the nature of "knowledge," and our epistemic

[43] A significant exception is Michael Williams (1996), who traces the problem to the assumption of foundationalism.

psychology more generally. Just as vision scientists investigate visual illusion to better understand how perception is successful, one way we can better understand the nature of knowledge is to investigate the causes of skeptical doubt.

In this essay I argue that (a) the skeptical problem has been misdescribed in two important ways. First, the linguistic tradition of skeptical diagnosis, including infallibilism, relevant alternatives theory, and linguistic contextualism,[44] all underestimate or ignore the role that the plausibility of a skeptical hypothesis plays in determining its salience to epistemic assessment. This is a mistake. I argue experimentally that folk knowledge attributions show that plausibility plays a key role in assessing the threat of a counter hypothesis. But because truly radical skeptical hypotheses are deeply implausible, any complete diagnosis of the skeptical threat must account for how our ordinary resistance to implausible error possibilities is overcome. Second, the linguistic tradition, particularly linguistic contextualism, also takes it that the intuition of conflict, between ordinary knowledge ascriptions and the conclusions of skeptical arguments, is clear, forceful and automatic. Defenders of contextualism of this sort, including Keith DeRose and Stewart Cohen, argue that the intuition of conflict is the product of a special kind of "semantic blindness." This too is false. Ordinary speakers do not automatically experience the intuition of conflict unless they are primed in particular ways. I argue experimentally that it is only highly reflective individuals, like philosophers, who experience the intuition of conflict without priming.

Thus, the conclusion of my experimental work is that skepticism is *not* intuitive, *nor* is it a natural by-product of our ordinary epistemic practices.[45] Radical skepticism is naturally accessible, but it is not natural. The skeptical problem, instead, only emerges for some people, in very special circumstances, as a by-product of our evolved epistemic psychology. In this essay I also argue (b) that the skeptical problem is best understood at the level of cognitive psychology. Describing the problem at this level gives a better account of both our ordinary indifference to implausible skeptical hypotheses as well as their occasional attraction. But this is an account of our epistemic psychology, and not knowledge in itself. I

[44] Not to mention infallibilists like Peter Unger and David Lewis (in the right conditions), invariantists like Timothy Williamson and John Hawthorne, and relativists like John Macfarlane respectively.

[45] The term "natural" deserves comment here; if by "natural" one means 'accessible,' then yes, in my terms skepticism is a natural by-product of our epistemic psychology; but, if by "natural" one means as the tradition has so long meant 'inevitable,' then by my lights, skepticism is *not* a natural by-product of constitution.

argue that understanding the skeptical problem can be accomplished at the psychological level and not at the normative level. This approach offers a much more promising diagnosis of the skeptical problem. It better explains the first personal character of the skeptical problem, and it does so without positing the implausible kinds of performance error that linguistic approaches must.

Before explaining how I'll defend this view, it's worth taking a moment to make clear exactly the problem I aim to describe and respond to. I am not trying to diagnose all forms of skeptical doubt. They are various, and range from underdetermination arguments to the argument from dispute. My focus is the ordinary argument from ignorance. I gave two examples of it above, but it has a straightforward form.

1. I don't know that not H
2. If I don't know that not H, then I don't know O
3. Therefore, I don't know O

Here H and O stand for any two incompatible hypotheses; O is a focal hypothesis, and H is a counter hypothesis that implies that O is false.. I call it the ordinary argument from ignorance, because I want to distinguish it from two philosophical variants that are related. One form of skepticism that begins from the possibility of error is high-standards skepticism. According to high-standards skepticism, genuine knowledge requires absolute certainty, and this is something we cannot achieve if we cannot rule out the radical skeptic's hypothesis. According to infallibilists like Peter Unger, while there is much that we can reasonably believe, there is little we can know with absolute certainly. High standards skepticism can get its start from the ordinary argument from ignorance, but as a complete position I think it involves further theoretical commitments, both about the nature of certainty and about reasonable belief, that are independent of AI. Ordinary AI can be used to motivate high-standards skepticism, but high standards skepticism is ultimately an interpretation of the result, not a consequence of AI without further assumption.

Likewise, strict skepticism is the view that skeptical arguments deprive our beliefs of any justification because the source of those beliefs, perception, does not entail that the beliefs are accurate. Like high-standards skepticism, I think AI can be used to motivate strict skepticism, but the conclusion of the argument on its own doesn't entail that all our beliefs are without justification. That conclusion only follows on further assumptions about the

relationship between perception and justification. My focus is just on our ordinary epistemic practices: our practices in attributing knowledge, and our practices in evaluating rival hypotheses. To see the difference, just notice that both high-standards skepticism and strict skepticism make universal claims about epistemic standards, certainty and justification, that don't appear in the argument, and cannot be connected to it without further assumptions. My focus is on our practices in evaluating particular claims. What more general principles they license is a different question I won't investigate here.

5. A Sketch of the Argument

Defending the position I outline here involves a few steps

Chapter II

My approach depends on experimental work. In this chapter I introduce experimental philosophy and the challenges it raises for epistemology. Experimental philosophy issues a prima facie challenge to the armchair diagnosis of the skeptical problem: a number of studies show that philosophers and ordinary people have different intuitions about many prominent epistemological thought experiments. If the skeptical challenge rests on intuitions, we can justifiably wonder if philosophers' intuitions are the same as the folk's.

Chapters III & IV

These chapters lay out a series of experimental results showing that the skeptical problem has been misdescribed. I argue that they show linguistic contextualism does not provide an adequate account of the skeptical problem. Chapter 3 shows that while contextualists are right that there is evidence that "know" is used in a context-dependent way, they're wrong to think that this phenomenon is directly responsible for the skeptical problem. I present evidence that ordinary speakers are both aware of contextually shifting standards, and that they are not prone to the kind of semantic blindness contextualists must posit to explain the skeptical problem. But if linguistic contextualism is wrong about semantic blindness, there are ways in to induce it. I present a further set of studies that show how it is possible to prime some subjects to have the intuition of conflict, and that this individual difference is better explained at the level of cognitive psychology . Chapter 4 argues that most speakers take the plausibility of a skeptical hypothesis into account when considering knowledge ascriptions. This runs counter to the predictions of linguistic

contextualism, infallibilism, and relevant alternatives theory, and many anti-skeptical invariantists. I show that a special kind of priming is required to motivate individuals to overcome their natural indifference to implausible counter hypotheses of the kind found in radical versions of AI.

Chapter V

This chapter lays out a positive, multi-process theory of the psychology of epistemic assessment. I argue that at least two (and perhaps more) distinct psychological systems are involved in epistemic assessment. The first is a permissive system that seeks evidence in support of a hypothesis. It is a fallibilist, justificatory system; it asks the question *can I believe P?* The second is a stringent system that seeks reasons to reject a hypothesis. It is an infallibilist system; it asks the question *must I believe P?* Contextual factors like the content of a target claim and its source can determine which system is active. The skeptical problem is a byproduct of a conflict between the outputs of these two systems. Arguments like AI can trigger the stringent system, for some people in some contexts, and this leads to the skeptical problem.

As a diagnosis of the skeptical problem, the multi-process theory I preview here has three virtues that previous linguistic approaches lack. First, it offers an account that is consistent with the folk ascription data of the *genuine* skeptical problem, as described in Chapters 3 and 4. Second, I show how this view can be integrated into a large literature in psychology on motivated reasoning and debiasing strategies, and how the skeptical problem is the logical extension of various tendencies in our folk psychology. Third, it offers a better account of the grip of the skeptical problem. Linguistic accounts cannot do justice to the 1st person *ambivalence*[46] of the skeptical problem. By their nature linguistic accounts draw their evidence from linguistic usage, and thus they draw upon our linguistic competence. They must also, however, posit pervasive linguistic incompetence to explain the grip of the skeptical problem – a kind of failure that persists even when, or despite, a mistake having been pointed out. It is both an unprecedented kind of performance error they must advert to, and a suspiciously local one. Consequently, their evidential base and their diagnosis

[46] By 'ambivalence' I mean the genuine, traditional sense, of being strongly pulled in two directions. Michael Williams describes this aspect of the skeptical problem as its bi-perspectivalism.

coexist in deep tension. Many psychological system conflicts, on the contrary, have just this kind of feature.

Moral puzzles, again, offer a very useful point of comparison. A number of leading moral psychologists (S. Nichols, J. Greene, & F. Cushman) argue that a number of different, domain specific systems exist to support moral decision making. Usually these systems have a proper domain, in which they operate and produce consistent outputs. However, certain special circumstances can activate both modules, producing inconsistent outputs, about which we have a similarly *ambivalent* response (the spur and bridge trolley problems, for example). The fact that there is no affective release from the dilemma is just a function of the outputs being generated by different systems with different intended functions. There can be no resolution that *turns off* the system that produces one of the conflicting intuitions. In this way therapy is not appropriate. There is no discovery that will reveal a correct result that also will *release* us from the tension. In this important way, skepticism is much more like a moral puzzle than a linguistic mistake. The multi-process process theory I propose does justice to the ambivalence of skepticism. But importantly, the kind of tension I describe here is only accessible when many points of resistance have been removed - so while skepticism is an *accessible* by-product of our epistemic psychology, it is neither *inevitable* nor *typical*.

This framework offers a naturalistic foundation for a diagnosis of the psychological grip of Cartesian skepticism. I argue that the biases and constraints in epistemic ascription that give rise to the skeptical problem only grip particularly reflective individuals, like philosophers and those who have been delicately primed for the skeptical argument. These biases can be understood as among the very many sub-personal, and automatic information processing events that occur prior to a representation becoming consciously available. As such, they tend to constrain, but not determine, our normative stance. This approach to the diagnosis of skeptical arguments explains both why they are gripping and unpersuasive.

Chapter 2

Experimental Philosophy & The Skeptical Problem

1. Introduction

The problem of radical skepticism has often been described as intuitive. Its intuitiveness is often taken as a sign that skepticism reveals a tension in our ordinary conception of knowledge. This is why skepticism also serves as the final test case for any theory of knowledge: it is a serious shortcoming of any theory that cannot explain how we have knowledge of the external world, or how the skeptic gets a grip on us. But recently experimental philosophy has seemed to show that philosophers' intuitions about some cases are quite different from ordinary people's intuitions about the same cases. It has also showed that intuitions can be highly unstable, and that small changes in presentation can have large effects on intuitions.

If the problem of skepticism is an intuitive one, but those intuitions are the armchair intuitions of philosophers, how can we be sure that their diagnosis of the problem really illuminates our ordinary conception of knowledge? And if the intuitions of philosophers are in some ways idiosyncratic, might not the diagnosis they offer also be idiosyncratic? The central claim I wish to advance in this chapter is this: folk intuitions about knowledge are essential to an accurate description of the most threatening versions of the skeptical problem. Consequently, an adequate response to the skeptical problem cannot be achieved through armchair diagnosis alone; instead it must begin with an empirically informed investigation of the component intuitions driving AI.

In this chapter I'll explore the challenge experimental philosophy raises for traditional diagnoses of the skeptical problem. I'll begin by surveying some of the uses to which intuitions have been put in contemporary epistemology, before laying out the empirical results that suggest ordinary intuitions about knowledge differ from philosophical intuitions about knowledge. I'll then look at a number of replies that have been offered in defense of philosophical appeals to intuition and show that whatever their promise in other areas of philosophy, they don't obviate the need for an empirical description of the skeptical problem. I'll conclude by considering a possible objection that experimental philosophy could level against the empirical approach to the skeptical problem I advocate.

2. The Method of Cases

Philosophy uses a number of methods. One part of the philosopher's toolkit at the center of attention and controversy over the last decade is the use of hypothetical thought

experiments and appeals to intuition about them. This method has been used across a number of branches of analytic philosophy.

One familiar instance of the method is Edmund Gettier's argument against the analysis of knowledge as justified true belief.[1] In his classic paper, Gettier argued the traditional analysis does not provide sufficient conditions for knowledge by presenting two counterexamples, both of which satisfied the conditions, but were intuitively not cases of knowledge.[2] What's distinctive about Gettier's argument is that it seems to rely on our ability to recognize clear cases of knowledge and non-knowledge as a means for testing a purported definition. Another prominent set of examples, from the long literature on proper names, are Saul Kripke's semantic arguments against the descriptive theory of names. In *Naming and Necessity* Kripke argued that knowledge of a uniquely identifying description is neither necessary (the Feynman case) nor sufficient (the Gödel case) for successful reference.[3] Kripke's arguments, like Gettier's, rely on the ability of ordinary speakers to recognize clear cases of successful reference and clear cases of unsuccessful reference.[4] A third well-known example is Philippa Foot's defense of the doctrine of double effect, which appealed to intuitions about the moral acceptability of diverting a runaway trolley from a crowded track to a less crowded track.[5] Foot, and the many subsequent elaborators of the trolley problem, rely on our ability to recognize, without the explicit use of theory, whether an action is intuitively permissible or not. According to the method, sometimes called the "the method of cases," the reader must render their own intuitive judgment about whether or not a hypothetical individual, S, in the hypothetical case, "knows P" or not, "refers to P" or not, has "acted permissibly" or not. Crucially, in each case the intuition is probative: the judgment about the hypothetical serves as evidence, for the purpose of argument, whether "S knows P," whether "S refers to P," and whether "S acts permissibly."

Using cases like Gettier's, Kripke's, or Foot's, as evidence for or against a favored position means accepting, in some way or other, that there are certain propositions that we can know just by thinking about them. Intuitive assessments like these play a similar role in

---

[1] Gettier, himself, doesn't use the term "intuition" in his classic paper, but see below for a detailed consideration of the now infamous cases.

[2] E. Gettier, 1963.

[3] S. Kripke, 1991.

[4] S. Kripke, 1991 (pg. 83 for the Gödel case; pg. 81 for Feynman case); Kripke, unlike Gettier, is explicit about his reliance on intuition (see pg.42) – despite the fact that the argument is for semantic externalism!

[5] P. Foot (1967).

motivating some important philosophical paradoxes. It seems intuitive that a pile of a million grains of sand is a heap. It also seems intuitive that after we've removed one grain we still have a heap. But, while each step seems intuitive, by easy stages we eventually arrive at the deeply counter-intuitive answer that one grain of sand is also a heap. Zeno's paradoxes of motion can be motivated as an inconsistent set of intuitions as well. Arrows can travel a given distance *d*. But before it can travel the full distance it must travel half the distance. Before it can travel half the distance, it must travel a quarter of the distance. Before it can travel a quarter of the distance, it must travel an eighth. And so on. So, an arrow can never travel a given distance *d*.

Likewise, the problem of radical skepticism appears to be an intuitive paradox as well. It can seem intuitive that we cannot rule out being deceived by Descartes' demon. And if we cannot rule it out, it seems that we don't know that we're not. But, if we don't know that we're not being deceived, we also don't know ordinary facts about the world, like that we have hands. So, it seems we don't know that we have hands. But, of course, we do know we have hands! As we've seen before arguments of this form are sometimes called the argument from ignorance (AI).

4. I don't know I'm not a BIV (brain-in-a-vat)
5. If I don't know I'm not a BIV, then I don't know I have hands
6. Therefore, I don't know I have hands

Keith DeRose, along with many others, has described the challenge of AI as an intuitive paradox, "the skeptical argument really is powerful…The argument is clearly valid…and each of its premises, considered on its own, enjoys a good deal of intuitive support."[6] It's worth taking a moment to recall exactly how the argument makes use of intuitions. The first premise seems to appeal to the intuition driving relevant alternative theory: if we cannot rule out a counter hypothesis, then we cannot know it's not the case. In chapter one I labeled this the *error intuition*. The second premise seems to rely on the intuitive principle that knowledge is closed under known implication: the principle that if I know that P, and I know that P implies Q, then I know that Q. In chapter one I called this the *closure intuition*. Now certainly this is not a belief that many non-philosophers have explicitly, rather the idea

[6] K. DeRose, 1999: 2-3. See, also, the preface.

is that they reason in accordance with it, which is why premise 2 seems intuitive. Our more general ability to recognize valid arguments explains why the conclusion follows. The final pair of intuitions are the *ordinary intuition*, that we do know we have hands, and the *intuition of conflict*. The first needs little comment: our knowledge of clear facts about our bodies is as well justified as any belief we might possess. The intuition of conflict is the immediate conflict we seem to see between the ordinary intuition that we have hands, and the conclusion of the skeptical which contradicts it. Together, these four commitments seem to be what commentators on skeptical arguments like AI are discussing when they consider its intuitive basis.

3. Intuition

But what is meant by the term "intuition" and why take intuitions seriously? These, as I'll explain, are questions of serious concern among analytic philosophers. Timothy Williamson has remarked that given the importance of intuitions in modern analytic philosophy, the absence of an agreed on account of what they are and the relationship between an intuition that P and the truth or falsity of P, is something of a "scandal."[7] Paul Boghossian has described invocations of intuitions as evidence as "the name for the mystery we are addressing, rather than a solution to it."[8] But before addressing these questions, we can start with Jennifer Nagel's broad characterization of intuitions as a label "for any immediate (or not explicitly inferential) assessment of any claim of interest."[9]

Appeals to intuition play an important role in analytic philosophy.[10] Appealing to intuitions as evidence has even been called philosophy's "standard justificatory procedure."[11] Saul Kripke offers a defense of appeals to intuition in *Naming and Necessity*,

> Some philosophers think something's having intuitive content is very inconclusive evidence in favor of it. I think it is very heavy evidence in favor

[7] T. Williamson (2007: 215).

[8] P. Boghossian (2000: 231).

[9] J. Nagel (2007: 793)

[10] Perhaps their contemporary use is an inheritance of analytic philosophy's past focus on conceptual analysis, as S. Laurence & E. Margolis (2003: pg. 277) describe it, "The distinctive feature of conceptual analysis is a reliance on people's intuitions about the application of concepts in actual and hypothetical circumstances." And even as philosophy's self conception has changed, the methods have remained.

[11] G. Bealer (1992), pgs. 99-138.

> of anything, myself. I really don't know, in a way, what more conclusive evidence one can have about anything, ultimately speaking.[12]

But intuition has been at the center of a number of different debates, which have been pursued along a number of different lines. It has been argued that intuitions are not a good source of evidence.[13] It has been argued intuitions are a good source of evidence.[14] It has been argued that intuitions are like beliefs.[15] It has been argued that intuitions are like perceptions.[16] It has been argued that intuitions are neither perceptual, nor doxastic, but a species of know-how.[17] It has been argued that intuitions vary from culture to culture.[18] It has been argued that intuitions do not vary from culture to culture.[19] It has been argued that philosophers are expert intuiters.[20] It has been argued that philosophers are not expert intuiters.[21] It has been argued that philosophical argument essentially relies on intuition.[22] It has been argued that philosophy does not rely on intuition at all.[23] It has been argued that intuitions are sensitive to presentation.[24] It has been argued that intuitions are not sensitive to presentation.[25]

I think that this kind of multi-decussate debate is in part a product of the difficulty in accurately characterizing intuition. Discussions of intuition tend to begin by identifying it in one of two related ways, either in opposition to conscious inferential thought, or in terms of a characteristic phenomenology. The first, more indirect, approach begins by describing the characteristics of everyday, conscious reasoning: it is under conscious control, it has a (more or less) logical structure, and it is temporally protracted. Intuitions are then identified

---

[12] S. Kripke,1991: 41-42.
[13] S. Stich, 1990, 2001; J. Hintikka, 1999; J. Alexander, R. Mallon & J. M. Weinberg, 2010; David Lewis, 1974.
[14] J. Knobe & S. Nichols, 2007; T. Williamson, 2004, 2007, 2011; G. Bealer, 1992, 1998, 2008; S. Kripke, 1991.
[15] A. I. Goldman & J. Pust, 1998; M. Deutsch, 2010; A. Gopnik & E. Schwitzgebel, 1998; J. Earlenbaugh & Bernard Molyneux, 2009; T. Williamson, 2004, 2007.
[16] J. Bengson (forthcoming), (unpublished dissertation); E. Chudnoff, 2011; G. Bealer, 1992, 1998, 2008.
[17] R. Neta (personal communication), presented at the 2012 NEH X-Phi institute, Tucson, AZ.
[18] J. M. Weinberg, S. Nichols & S. Stich, 2001; E. Machery, R. Mallon, S. Nichols & S. Stich, 2004.
[19] J. Nagel, 2012; B. Lam, 2010.
[20] S. D. Hales, 2006; K. Ludwig, 2007; S. Matthew Liao, 2008.
[21] J. M. Weinberg, C. Gonnerman, C. Buckner & J. Alexander, 2010.
[22] S. Stich, 2001.
[23] H. Cappelen, 2012.
[24] S. Cullen, 2010.
[25] F. Jackson, 1998.

as mental events that are not like *that*: they are automatic, they don't have a logical structure, and they are fast. Here's and example of such an attempt:

> Intuitions are odd critters: intellectual happenings in which it seems to us that something is the case without arising from our inferring it from any reasons that it is so, or our sensorily perceiving that it is so, or our having a sense of remembering that it is so.[26]

The second, phenomenological approach begins by directly pointing out that some thinking involves a species of thought which is fast, spontaneous, and often particularly clear or forceful. Intuitions, it is claimed, are just *those* states. In the following I want to be as ecumenical as possible. I think it's more than plausible that "intuition" is more of a genus term than a species term. I think it's quite plausible that many different psychological systems are involved in generating fast, automatic, non-inferential, non-perceptual assessments. The major contrast that I want to mark is the one between consciously directed assessments and fast automatic assessments. By "intuition" I'm pointing at the latter. But one disambiguation is worth pursuing before we go on. Sometimes we use "intuition" to mark a proposition's having intuitive content, i.e. the proposition "just makes sense," i.e. that it seems obviously true. Perhaps the sentence "bachelors are un-married men" is one such proposition. Other propositions don't make sense, for example "I know it's raining but I don't believe it." Other times we're marking that an assessment has been made quickly, but it doesn't follow that the content of the assessment is easy to understand, or "just makes sense." So, perhaps a commercial pilot might have the automatic thought, "increase the trim" but the reasons for this are complex and depend on expert training. This is not the kind of thought just anyone would have, or which just makes prima facie sense - I don't even know what "trim" is, except you adjust it in planes. And then, of course, these two features can coincide. Simple mathematics is one such place, "2+2=4" is something which has intuitive content, and which we can assess in a fast and automatic way. By in large my focus is on the latter sense of "intuition," of it being fast and automatic, but without committing one way or the other on whether the content of the intuition is in some way seemingly true. Of course that something has intuitive content often coincides with it being the product of fast, automatic cognition. This is an important part of the force we invest in

[26] J. Weinberg, 2007: 318.

them. But many cognitive biases are the result of fast automatic cognition, such as the gambler's fallacy, and individual instances do not have intuitive content.

Our intuitions can be very powerful. Confronted with the choice between allowing a runaway trolley to hit five workmen, and diverting it onto a track with just one, most people are immediately struck by what they should choose. This forcefulness is no doubt part of the reason why appeal to intuition and the method of cases has been such a dialectically effective method of argument. It certainly goes a long way toward explaining the prominence of thought experiments and the method of cases in philosophy.

But the rather obscure origins of intuitions raises a number of questions. Are they good evidence? Why are they so psychologically gripping? And, are they universally shared? A large part of the recent interest in intuitions and the method of cases centers on just these kinds of questions. This shouldn't be surprising, for without persuasive answers to any of these questions, it is very hard to offer a convincing justification of treating intuitions as probative. We can quite rightly ask why we should privilege the intuition in the Gettier case that Smith *doesn't* know who'll get the job over the implication of the theory that Smith *does* know? Absent any answers to these questions it becomes hard to explain what makes intuition better than instinct, prejudice, or mere guesswork.

Experimental philosophy has emerged in the last decade in response to worries like these. Its most prominent early exponents include Stephen Stich, Shaun Nichols, and Jonathan Weinberg. Their first concern was that appeals to intuitions about critical cases were essential to the standard justificatory procedure in philosophy,[27] and that those intuitions were not as widely shared as the traditional method presumed.[28] Some of the most discussed results of the early literature were cross-cultural variation in intuitions, variation between individuals with different socio-economic statuses, and lack of consensus about cases even among relatively narrowly defined groups of participants.[29]

---

[27] George Bealer (1992) first coins the term "standard justificatory procedure," but it is adopted by others, including Hilary Kornblith (2004), Joel Pust (2000), and Timothy Williamson (2007: 215).

[28] Stich first aired this concern in *The Fragmentation of Reason*, and followed it up experimentally, as we'll see, in 2001.

[29] For the origin of concerns about cultural differences, see: R. Nisbett et al. (2005); for a work that was enormously influential in spurring renewed interest in cross-cultural work see Henrich et al. (2001); and for a much more recent, and broader attack, on the narrow cultural base upon which global inferences about human nature have been founded in the social sciences, see J. Henrich et al. (2010). For worries about SES and intuitive responses, see J. Haidt (1996). More recently still, as the social sciences and experimental philosophy have become much more concerned with replicating results, some of these findings have come under pressure.

In fact, there are a number of different worries one might have, but they are summed up in a general way by Stacey Swain, Josh Alexander, and Jonathan Weinberg when they write:

> [There is] an existing body of empirical research demonstrating that intuitions vary according to factors irrelevant to the [philosophical] issues... [This research] raises questions about the reliance on intuitions…We take the growing body of empirical data impugning various intuitions to present a real challenge for philosophers who wish to rely on intuitions as evidence.[30]

This challenge is sometimes called *restrictionism*. The animating claim of restrictionism is that the use of intuitions as evidence should be curtailed. How much? That depends: some versions of the restrictionist challenge are more revisionist than others. The strictest versions suggest that there is little to no place for the use of intuitions in philosophy.[31] Weaker versions of restrictionism think that the only problem with philosophical appeal to intuition is that philosophers shouldn't look merely at their own intuitions.[32] In other words, that the method is sound but that appeals to intuitions are experiments with a sample of N=1.

The exact nature of the restrictionist challenge depends on how one construes the object of inquiry. For instance, some epistemologists, like Alvin Goldman, explicitly take themselves to be studying the *concept* of knowledge as a psychological structure.[33] Others, like Timothy Williamson, take themselves to be investigating the *nature* of knowledge, and not a psychological structure.[34] To begin to explore these challenges and others that emerge

---

See Brian Nosek and the Open Science Foundation (https://osf.io) for general worries and methods of replication. See Joshua Knobe and the Experimental Philosophy Replications Archive (http://goo.gl/LAiGrz) for more information on these studies in particular.

[30] Swain, Alexander, & Weinberg, 2008: 153.

[31] S. Stich, 2001, and J. Weinberg, 2007.

[32] See J. Knobe & S. Nichols, 2007. Here's how they characterize the positive project: "a critical method for figuring out how human beings think is to go out and actually run systematic empirical studies. Hence, experimental philosophers proceed by conducting experimental investigations of the psychological processes underlying people's intuitions about central philosophical issues." (pg. 3) And again, on how the method is applied: "As far as we know, no experimental philosopher has ever offered an analysis of one concept in terms of another. Instead, the aim is usually to provide an account of the factors that influence applications of a concept, and in particular, the internal psychological processes that underlie such applications. Progress here is measured not in terms of the precision with which one can characterize the actual patterns of people's intuitions but in terms of the degree to which one can achieve explanatory depth" (pg. 5).

[33] See A. Goldman, 2007, 2010. Recently Goldman (personal communication) suggests that the experimental data to be discussed is evidence that the verb "know" is polysemous.

[34] T. Williamson says (2007: 206) "the primary concern of epistemology is with the nature of knowledge, not with the nature of the concept of knowledge." For discussion of the point, see S. Stich: (2014: 3 - online early view mnscpt.). See also J. Nagel (2012: 495).

from the experimental philosophy movement, as well as the possible responses to them, I'll first look at some of the seminal results that initiated this critique.

4. Thought Experiments and Actual Experiments

To date there has been little direct experimental work on the skeptical problem. The restrictionist challenge began by focusing on cases used to develop a positive account of knowledge, and I'll follow that cue in explaining the approach. First I'll look at Gettier's original case-based argument against the justified true belief (JTB) theory of knowledge, and subsequent attempts to test his conclusions empirically. I'll then consider the central restrictionist challenge to the method of cases, which is based on these results, and which I'll call the *general variation argument*. Second I'll look at some problems with these early experiments and how more recent work on Gettier cases helps address them. Third, I'll turn toward the skeptical problem more fully, and I'll look at the results of three early experiments, focusing respectively on Dretske's zebra cases, the brain-in-a-vat thought experiment, and salience effects, all of which suggest that philosophers and non-philosophers may differ when it comes to the intuitions which drive the skeptical problem. These experiments have a number of problems as well, and I'll outline an agenda for rectifying them - an agenda I'll take up in Chapters 3 and 4. Finally, I'll consider three of the most important objections to experimental philosophy and the restrictionist challenge, and I'll argue that they do not obviate the need for an empirical description of the skeptical problem.

4.1 Gettier Cases

To begin to get a sense of how experimental philosophy initiated its challenge to analytic epistemology, consider the familiar examples from "Is Justified True Belief Knowledge?". In it Edmund Gettier sets out an objection to the traditional analysis of knowledge, according to which an agent S knows that P *iff* i) P is true , ii) S believes P, and iii) S has justification for believing that P (in some way).[35] The structure of Gettier's

[35] Gettier (1963) considers formulations drawn from Plato's *Theaetetus*, Roderick Chisholm's *Perceiving: A philosophical study* (1957), and from A. J. Ayer's *The Problem of Knowledge* (1956), according to which the justification condition is thought to be fallibilistic: i.e. it can be satisfied without thereby guaranteeing the truth of the proposition. This condition has been formulated in a number of ways; Gettier cites Chisholm as requiring one to have adequate evidence, and Ayer as requiring the right to be sure (pg. 121).

argument against this analysis is straightforward. He aims to show that the analysis is faulty by showing that satisfaction of the conditions of the analysis is not sufficient to guarantee the agent knowledge. He begins by introducing the following case:

> Suppose that Smith and Jones have applied for a certain job. And suppose that Smith has strong evidence for the following conjunctive proposition: (d) Jones is the man who will get the job, and Jones has ten coins in his pocket. Smith's evidence for (d) might be that the president of the company assured him that Jones would in the end be selected, and that he, Smith, had counted the coins in Jones's pocket ten minutes ago. Proposition (d) entails: (e) The man who will get the job has ten coins in his pocket. Let us suppose that Smith sees the entailment from (d) to (e), and accepts (e) on the grounds of (d), for which he has strong evidence. In this case, Smith is clearly justified in believing that (e) is true. But imagine, further, that unknown to Smith, he himself, not Jones, will get the job. And, also, unknown to Smith, he himself has ten coins in his pocket. Proposition (e) is then true, though proposition (d), from which Smith inferred (e), is false. In our example, then, all of the following are true: (i) (e) is true, (ii) Smith believes that (e) is true, and (iii) Smith is justified in believing that (e) is true. But it is equally clear that Smith does not know that (e) is true.[36]

Gettier's argument against the traditional analysis of knowledge is both fascinating and important. He first describes a case in which an agent, Smith, has a belief that satisfies the three target conditions in the analysis of knowledge, but which is *intuitively* not a case of knowledge. This is the critical step: Gettier invites the reader to share his own immediate classification of the case as *not counting as knowledge.* Gettier himself does not use the term "intuition," nor does he make any explicit remarks about his method or how we are to evaluate the case; all he offers is that "it is equally clear that Smith does not know." There are really only two places in evaluating the example where the reader is invited to render a judgment. Gettier stipulates that Smith believes the proposition (e), and that (e) is true, and thereby stipulates that conditions (i) and (ii) are satisfied. He then states that Smith competently deduces (e) from a false, although well supported, proposition (d). Here we are invited to make a judgment. Gettier takes competent deduction from an evidentially well-supported proposition to amount to justification. Allowing fallibilism, this is plausible, if contestable.[37] More importantly, there are theories of justification that can be applied in

[36] *Ibid*, pg. 122

[37] It, of course, is not exactly news that Gettier's paper has inspired a large literature, including attempts to offer (i) necessary and sufficient conditions on knowledge, (ii) attempts to explain why (i) is impossible, and (iii) attempts to see if the judgments he invites us to share are in fact shared by individuals outside of philosophy.

assessing the third condition.[38] The only other place where the reader is asked or positioned to render a judgment concerns whether Smith *knows* that (e). But absent a theory of knowledge with which to evaluate the case, since that's just what the analysis itself aims to be, the reader must simply render a judgment about whether Smith *knows*. They must do so without consciously reflecting on a theory of knowledge. They must do so intuitively.

Understanding the response Gettier invites his readers to share in completing the argument deserves some scrutiny. It has become commonplace to describe the reader's reaction to the question whether or not "Smith knows (e)" as an intuition. As I said, Gettier himself never does. Nor does this way of framing the problem emerge immediately. Jaakko Hintikka speculates that the use of the method of cases in philosophy, where the critical classification turns on an intuition, was inspired by the success of Chomskyan linguistics.[39] The first reference I can find that explicitly describes the Gettier method in something like the form of what I've called the method of cases is in Gilbert Harman's 1968 "Knowledge, Inference, and Explanation."[40] There he explicitly describes his method as one of intuitively testing analyses of "knowledge" against cases in a defeasible way.

4.2 Gettier Experiments

Gettier himself doesn't propose an analysis of knowledge, but many who followed him did. If one thinks that the output of this kind of investigation is something with

---

But, for a review of responses of type (i), and various debates surrounding whether Gettier's cases are actually justified belief, or if a fourth condition on knowledge is required, see Robert Shope's *The Analysis of Knowing* (1983), especially sections 1.4 and 1.5, pgs: 21 - 33. Indeed, to return to the satisfaction of the third condition, the judgment Gettier renders is consistent with the tradition of analysis he is addressing, for Ayer says of his third condition, that the right to be sure "can be earned in many ways" (1956: 25).

[38] Though, of course, the fundamental paradox of analysis which I am about to raise confronting "knowledge", or any concept, can also raised concerning "justification."

[39] J. Hintikka, 1999: 127.

[40] I conducted a search with Google "Ngram," and this is the first paper using both "Gettier" and "intuition." In it Harman is attempting what he calls, "an *empiricist analysis of knowledge*" (emphasis original). He says of his method: "I take an analysis to be any interesting set of necessary and sufficient conditions. Although I shall not offer an analysis of the meaning of "know" (whatever that would be), I shall appeal to your intuitions about hypothetical cases. I shall claim, for example, that a person can come to know something when he is told it, or when he reads it in the newspaper. Although I may seem to appeal to what one would ordinarily say about such cases and for this reason may seem to be doing "linguistic analysis," I am interested in what is true about such cases and not just in what we say about such cases. But, since I want to test the assumption that ordinary judgments about knowledge are usually correct, trust your natural inclinations about the cases I describe. Consider what you would naturally and ordinarily judge, if you were not doing philosophy. Fine distinctions made in ordinary judgments become blurred when these judgments are made in philosophical contexts." (pg. 164) Though Harman does describe something here like the method of cases, it's important to emphasize that he doesn't take accord with intuitive categorization judgments to be the *only* test of an analysis or theory.

normative force, then it seems like the project assumes that these intuitions are widely shared. But Jonathan Weinberg, Shaun Nichols, and Stephen Stich (WNS)[41] worry that if people have different epistemic intuitions about the very same cases, then we have no answer to the question "why should we privilege our own intuitions rather than the intuitions of some other group?"[42] Stich first airs the possibility in *The Fragmentation of Reason*; in "Normativity and Epistemic Intuitions" he and his collaborators set out to test two hypotheses: that epistemic intuitions vary from culture to culture, and that epistemic intuitions vary from socioeconomic group to socioeconomic group. The inspiration for WNS's cultural hypothesis was work done by Ara Noranzayan and Richard Nisbett suggesting that Westerners are more likely to focus on causal concerns in analyzing a situation, while East Asians are more likely to focus on similarity.[43] Previous work by Jonathan Haidt showing that socio-economic status (SES) is a better predictor of some moral intuitions than culture of origin prompted WNS to investigate it as well.[44] Neither of these investigations suggests a specific hypothesis about epistemic intuitions. Rather, according to WNS their interest is that they raise the general possibility that intuitions about *normative* subject matters vary across these groups.

WNS conduct nine different experiments using a number of familiar philosophical thought experiments to test for cultural and socio-economic variation in intuitions. The experiments include: three cross-cultural comparisons of variations on Keith Lehrer's "Truetemp" cases;[45] two cross-cultural variations of Gettier's cases; two variations of a "conspiracy" case of their own devising, one involving a cross-cultural comparison and another involving a SES comparison; and finally two versions of Fred Dretske's "Zebra" case, one involving a cross-cultural comparison and another involving a SES comparison. To explain their results and their import, I'll focus on their Gettier and Zebra results.

In their first experiment WNS tested their cultural variation hypothesis by presenting a Gettier-style vignette to Rutgers undergraduates who self-identified as having Western

[41] Because many of the early papers in the literature involve the same authors, rather than abbreviate attributions, e.g. Weinberg et al., I'll use acronyms, e.g. WNS.
[42] Weinberg, Nichols, & Stich (2001: 9 mnscpt).
[43] A. Noranzayan et al., 1999; A. Noranzayan, 2000; R. Nisbett et al., 2001; A. Noranzayan et al., 2002; R. Nisbett et al., 2005.
[44] J. Haidt et al., 1993.
[45] K. Lehrer, 1990.

(*n*=66), East Asian (*n*=23), or South Asian (*n*=23) cultural backgrounds. Each group was presented with the following, simplified version of a Gettier-case.

> Bob has a friend, Jill, who has driven a Buick for many years. Bob therefore thinks that Jill drives an American car. He is not aware, however, that her Buick has recently been stolen, and he is also not aware that Jill has replaced it with a Pontiac, which is a different kind of American car. Does Bob really know that Jill drives an American car, or does he only believe it?
>
> REALLY KNOWS ONLY BELIEVES

They found that while 74% (49/66) of Western participants shared the typical Gettier intuition and responded that Bob "only believes," only 43% (10/23) of East Asian and 39% (9/23) of South Asian participants gave the same response. The difference is statistically significant.[46, 47]

---

[46] As literally the first piece of experimental work done by philosophers, this study does have some understandable methodological faults. First, they do not systematically report any demographic data including sex or age, beyond saying that the participants were drawn from "lower-division lectures at Rutgers" (mnscprt. pg. 20). They don't report their methods, so it's unclear how many vignettes participants encountered, in what order, and in what setting. Second, the same data-set (Westerners) is used for multiple comparisons, but it is unclear whether the data-sets to which it is compared (East Asians or South Asians) were collected at the same time, or in the same class. Third, they don't report an effect size (though it can be recovered because they do report their raw data). Fourth, WNS choose a binary forced-choice design for collecting their data. Participants are asked to choose between "knows" or "only believes." They do not, however, discuss why they choose this design. Forced-choice designs do not measure confidence very well: which is to say they have difficulty distinguishing a bell-curve distribution closely clustered around the midpoint that can be associated with guessing among a population that is not very confident, a bi-modal distribution where all responses are at the poles, or something in between. As such, including an interval scale question about the respondent's confidence would have made interpreting the data easier. See the following note for a reanalysis. Fifth, and most importantly, they do not run controls. Their comparison is between group's responses to the same vignette. While this does allow them to validly conclude that there are statistical differences between the groups, it does not give them any information on what those differences are, or if they have any connection to the philosophically significant point. A better design would look at independent cases, one that *doesn't introduce* the Gettier error, and one that *does*; this would allow them to make conclusions about the philosophical claim about the case. Otherwise, it could merely be facts about cars or other irrelevant details that the groups are responding to. See Chapter 4 for improved designs in experiments I've conducted.

[47] WNS don't report effect sizes for any of their studies. They do, however, the report raw data, from which it is possible to determine this ourselves. They also report their results as two comparisons, and test these comparisons using a 2x2 contingency table. This isn't appropriate in this case, because they use the same sample of results from Westerners in both of their cross-cultural comparisons. So, comparing Western responses to East Asians, they report p=.0064 using Fisher's Exact test, and comparing the same sample of Western responses to South Asians, they report p=.002 using the same test. A more appropriate test would be a 3x2 contingency table and the Chi-Square statistic. Using it, and rerunning their raw data, there is an effect on knowledge attribution according to culture, where X^2=12.42 (2), p = .002, and Cramer's v = .333, indicating a "moderate" effect. To say more about the significance of the pair-wise comparisons it would be most appropriate to use Bonferonni's correction and halve the significance value for each comparison.

In their second set of experiments they test the SES hypothesis by comparing the epistemic intuitions of groups distinguished by years of education; participants were coded as low SES if they self-reported as never having attended college, and high SES if they self-reported as having attended college at some point. They presented a variation of Fred Dretske's fake Zebra case to (n=24) low SES participants and (n=34) high SES participants. Each group was presented with the following probe.

> Pat is at the zoo with his son, and when they come to the zebra cage, Pat points to the animal and says, "that's a zebra." Pat is right — it is a zebra. However, given the distance the spectators are from the cage, Pat would not be able to tell the difference between a real zebra and a mule that is cleverly disguised to look like a zebra. And if the animal had really been a cleverly disguised mule, Pat still would have thought that it was a zebra. Does Pat really know that the animal is a zebra, or does he only believe that it is?
>
> REALLY KNOWS ONLY BELIEVES

They found that 33% (8/24) of low SES participants answered that Pat "really knows" that the animal is a zebra, while only 12% (4/34) of high SES participants answered similarly. This result is statistically significant.[48]

4.3 The Variation Argument

All in all, WNS report nine different studies that support the claim that intuitions about epistemological cases vary across culture and socio-economic status. WNS argue that their results pose a significant challenge to analytic epistemology. They issue the challenge by way of what I call "the variation argument."

1. Intuitions are evidence for normative theories
2. If intuitions vary by group, then norms will vary by group
3. Intuitions vary by group
4. Norms do not vary by group
5. Therefore, intuitions are not evidence for normative theories.

[48] Again, WNS don't report an effect size. Using their reported raw data to calculate one results in an (odds ratio) $\sigma = 3.75$. In other words, high SES participants are nearly 4 times more likely to share the Zebra intuition than low SES participants, on the assumption that what varies between groups *is* the error intuition. See note 47 concerning controls.

They defend their claim in premises 1 and 2 by offering Alvin Goldman as a paradigm example of an analytic epistemologist. They quote Goldman when he says, we aim for the theory of justification that is "embraced by everyday thought and language"[49] and that such a theory is "supported to the extent that implied judgments accord with such intuitions, and weakened to the extent that they do not."[50] They defend 3 by pointing to their empirical work on Gettier intuitions and Dretske-style error cases. As such, they set up a dilemma for analytic epistemology: it must either give up appealing to intuitions as data, or it must embrace a serious form of relativism that allows norms, such as those that govern knowledge and justification, to vary by group. WNS argue, further, that it is a particularly uncomfortable form of relativism to embrace, for given their results concerning SES, anyone opting for this road would have to adopt one epistemology for rich people, and a different one for poor people.[51] They conclude that traditional epistemology, as it relies on appeals to intuitions,

> seems a rather bizarre way to determine the correct epistemic norms. For it is difficult to see why a process that relies heavily on epistemic intuitions that are local to one's own cultural and socioeconomic group would lead to genuinely normative conclusions.[52]

But what explains the difference in intuitions? WNS don't endorse any particular explanation, but they offer two intriguing speculations.

> It has often seemed to us that students' epistemic intuitions change as they take more philosophy courses, and we have often suspected that we and our colleagues were, in effect, teaching neophyte philosophers to have intuitions that are in line with those of more senior members of the profession. Or perhaps we are not modifying intuitions at all but simply weeding out students whose intuitions are not mainstream. If either of these is the case, then the intuitions that "we" use in our philosophical work are not those of the man and woman in the street, but those of a highly trained and self–selecting community.[53]

---

[49] A. Goldman, 1986: 58; quoted in WNS, 2001: 7.

[50] A. Goldman, 1986: 66; also quoted in WNS, 2001. Weinberg et al. (2001: 7 mnscpt) note that Goldman has responded to their studies by drawing attention to later work of his, according to which intuitions are evidence for our concrete psychological concepts, and as such philosophers should be worried about drawing their evidence solely from a "fairly homogenous subculture" (Goldman, 1992: 160).

[51] WNS, 2001: 27 – 28 mnscpt.

[52] WNS, 2001: 36 mnscpt.

[53] WNS, 2001: 17, mnscpt.

If training, or selection bias, really does result in philosophers having different intuitions about epistemological thought experiments, then it's open to us to wonder whether the intuitions behind the skeptical problem are widely shared beyond philosophy. But we shouldn't jump to conclusions. On the one hand there is ample evidence that training can affect intuitive assessment. As we saw in chapter one, most people have the Wason selection task intuition, but expert training helps philosophers avoid it.[54] This kind of consideration suggests that perhaps philosophers' intuitions about the skeptical problem *will* be different than non-philosophers'. On the other hand, there is ample evidence that some intuitions are deeply resistant to even expert training. Daniel Kahneman and Amos Tversky's work on statistical reasoning is an example of this. Most people have difficulty reasoning about probability, and some of the errors they make are systematic. We tend to ignore the base-rate of a phenomenon when making statistical judgments, and we sometimes think that the conjunction of two events is more likely than either of the conjuncts individually (the Linda the bank teller fallacy). Kahneman and Tversky found that even graduate students in statistics at Stanford are nearly as prone to these intuitive biases as non-statisticians.[55] So, while we have reason to worry philosophers have different intuitions, we also have reason to look further.

4.4 Experimental Control and Generality in Gettier Experiments

These initial experiments spurred the growth of experimental philosophy.[56] Since their initial appearance, these studies and the general variation argument, and the claim that intuitions vary by group have all been challenged in a number of ways. I'll take up the most important of these challenges momentarily. But before doing so, I'll address two prima facie worries about WNS's conclusions.

---

[54] Genuinely *expert* training matters; see Chapter 5 for a discussion of how hard it is to change performance on the selection task through training.

[55] D. Kahneman & A. Tverskty, 1983: 300.

[56] These were not the first attempts to empirically test philosophical claims. Lewis Petrinovich first began studying folk responses to trolley problems in the early '90's; John Mikhail independently began investigating trolley problems empirically around 2000, although he did not publish any results until 2007. Arne Naess is credited by Anthony Appiah (APA Presidential Address, 2009) as pursuing the first piece of experimental philosophy in the 1930's by using subjective survey methods to assess lay theories of truth; see his *'Truth' as Conceived by Those who are Not Professional Philosophers* (1938). Of course, there is a much older tradition according to which what we now call "science" was called "natural philosophy"; indeed, this is how Newton described his researches in the *General Scholium.*

An initial worry is that one study cannot show that all intuitions about Gettier cases vary. To be clear, the inferential statistic they apply allows them to infer that in the broader population the likelihood that the pattern of responses *to those experimental probes* is random variation, is less than 2 in a thousand (and less than 3 in a hundred in their Zebra case). In both cases this is below the conventional threshold of significance in the psychological sciences. So, in a word, these experiments allow you to conclude that people will respond in the same way to the same prompts. They license only a much weaker generalization about Gettier cases per se. Before concluding that folk and philosophical intuitions about Gettier cases really are different, it's important to establish that the result is general, and depends on what we can call the "gettierizing feature," (luck, or a lack of a causal connection, or whatever it is that causes philosophers to deny knowledge to the vignette's victims) and not an idiosyncratic feature of WNS's vignettes. To do so, it's necessary to run further tests of similarly classed Gettier cases, and see if the pattern of responses is the same. This introduces a kind of chicken and egg question. There is no consensus on what makes a Gettier case a Gettier case. Some accounts opt for the fact that "luck" of some kind is involved, others opt for the lack of a causal connection, while others opt for a false lemma in the inference.[57] In fact, the identity criteria for Gettier cases are as numerous as the analyses of knowledge that have been offered in response.

A second, and very simple worry, which has not been pressed in print to my knowledge, is that WNS do not run control studies in their experiments. Their goal is, as stated, to compare responses between groups. They conclude that Gettier intuitions vary across groups because responses to their question do. But this conclusion isn't really licensed by the experiments they run. In the Gettier case described above, 74% of Western participants say that the agent in the vignette "only believes" that "Jill drives an American car" whereas only 39% of South Asians share this Gettier intuition. But what really is going on is that 39% of South Asians share this *response* to this vignette. We cannot yet infer that the difference in their responses is caused by the feature that philosophers agree makes this a Gettier case in the first place. It could be that the difference is driven by the subject matter, or the names of the agents in the vignette, or the number of words in the vignette. Perhaps

[57] See William Lycan (2006) for a contemporary defense of the false lemma position, and for a survey of the many current approaches to the problem, all leading to a diagnosis of the failure to arrive at a satisfactory analysis.

what varies between groups is just beliefs about what can be known about cars, or about what can be known by people named Bob.

It is true that the Gettier feature is the likeliest candidate, as the responses rendered by the participants do concern knowledge and belief, and there is no prior reason to think that a car's make and model, a person's name, or the length of a story are at all relevant to epistemic evaluations. Nevertheless, a better design would have presented minimal pairs to participants across cultures: one vignette which involved the putative Gettier feature, and one which lacked it. Then it would be possible to see if the same *difference* in responses to the paired cases appeared across culture, SES, or level of philosophical training. For instance, working from their vignette, a good control might be the following:

Control Case	Gettier Case
Bob has a friend, Jill, who has driven a Buick for many years. Bob therefore thinks that Jill drives an American car. He is not aware, however, that her Buick has recently been **painted brown**, and he is also not aware that Jill **also owns a Pontiac**, which is a different kind of American car.	Bob has a friend, Jill, who has driven a Buick for many years. Bob therefore thinks that Jill drives an American car. He is not aware, however, that her Buick has recently been **stolen**, and he is also not aware that Jill **has replaced it with** a Pontiac, which is a different kind of American car.
Does Bob really know that Jill drives an American car, or does he only believe it?	Does Bob really know that Jill drives an American car, or does he only believe it?

Table 1.

The simple case is, seemingly, a straightforward case of justified true belief. The second is, seemingly, a straightforward case of a "luck" Gettier case. By comparing responses between groups to these two cases, it's possible to see if they are equally responsive to the putative Gettier feature. If, per the implausible hypothesis I aired earlier, South Asian participants were really responding to the names of the agents, or to the makes of the cars, then the presence or absence of the putative Gettier feature would not affect their responses, and they would respond in the same way to both vignettes.

But this raises an important question. Does this methodological failing undermine the evidence for premise 3 of the general variation argument? It may seem to. For, if the failure to run controls undermines the conclusion that what varies between groups are Gettier intuitions, then it is less clear that anything of philosophical significance varies between groups. It could still be argued that epistemic norms vary by group, if it turns out

that beliefs about what can be known by people named "Bob" vary by group. But this could be explained away as a simple false belief. Many groups have had many false beliefs, about the shape of the earth, the nature of the stars, among much else. What is truly threatening about WNS's challenge is that it attacks the *methods* by which philosophers refine folk theories of knowledge in formulating normative theories. But what their tests might simply show are differences in folk theories, and this is much less threating.

Perhaps WNS's challenge could be rehabilitated on its own grounds in the following way. Part of the outstanding difficulty of developing an analysis of knowledge in response to the Gettier problem is that the structure of our concept of knowledge and the reasons for our response to Gettier cases are opaque. All anyone has is their own intuition about the case, and it is not pellucid: what causes the intuition is not part of the content of the intuition. As such, it would be open to anyone who wished to press WNS's line to argue that *philosopher's* thought experiments make the same kind of control-case mistake. It's possible that what is really going on in standard presentations of the Gettier case, just to air a few possibilities, is a sophisticated sort of priming, or argument from authority. Moreover, because N=1, it is not possible to run independent controls, and consequently it is much more difficult for philosophers to introspectively determine that previous exposure to the control case before considering the target case, or vice versa, hasn't affected their naïve intuition.

4.5 Recent Gettier Experiments

Recently, a number of methodologically much more sophisticated tests of the Gettier case have been run, which allay many of the concerns addressed in section 2.4, and further support a version of the general variation argument.[58] Christina Starmans and Ori Friedman conducted a number of well-controlled tests, of a variety of conditions, which they argue support the claim that the folk conception of knowledge is accurately described by the JTB

[58] Even more recently, in forthcoming, but unpublished work, Edouard Machery has run a series of studies on cross-cultural differences in Gettier intuitions as part of project on intellectual humility co-directed with Stephen Stich. Among many of their results, they found cross-cultural variation in Gettier intuitions between groups with American, Brazilian, Central African, and Japanese Backgrounds (personal communication & results presented at Epistemology for the Rest of the World, Tokyo, Japan, August 2013). These results would, if they hold, license a stronger version of the variation argument than Starmans and Friedmans' results.

theory and that the folk do not share the philosophers' Gettier intuition. The central cases they use are below, in Table 2. Text that is different between cases is bolded.[59]

Control Case	Gettier Case	False Belief Case
Peter is in his locked apartment reading, and is about to have a shower. He puts his book down on the coffee table, and takes off his black plastic watch and leaves it on the coffee table. Then he goes into the bathroom.	Peter is in his locked apartment reading, and is about to have a shower. He puts his book down on the coffee table, and takes off his black plastic watch and leaves it on the coffee table. Then he goes into the bathroom.	Peter is in his locked apartment reading, and is about to have a shower. He puts his book down on the coffee table, and takes off his black plastic watch and leaves it on the coffee table. Then he goes into the bathroom.
As Peter's shower begins, a burglar silently breaks into the apartment. The burglar takes Peter's black plastic watch, replaces it with **an identical black plastic watch**, and then leaves. Peter is still in the shower, and did not hear anything.	As Peter's shower begins, a burglar silently breaks into the apartment. The burglar takes Peter's black plastic watch, replaces it with **an identical black plastic watch**, and then leaves. Peter is still in the shower, and did not hear anything.	As Peter's shower begins, a burglar silently breaks into the apartment. The burglar takes Peter's black plastic watch, replaces it with **a dollar bill**, and then leaves. Peter is still in the shower, and did not hear anything.
Peter _____ that there is a **book** on the table. (Really knows / Only thinks)	Peter _____ that there is a **watch** on the table. (Really knows / Only thinks)	Peter _____ that there is a **watch** on the table. (Really knows / Only thinks)

Table 2.

They found that approximately 90% of participants attributed knowledge to the agent, Peter, in the control case, and that 70% percent attributed knowledge in the Gettier case. Only 10% attributed knowledge in the false belief case.[60] They conclude that participant's folk

[59] C. Starmans & O. Friedman, 2012: 274.

[60] In each case, Control, Gettier, and False Belief, participants' responses to the dichotomous knowledge question were significantly different from chance (binomial tests, all *p*s <.003); there was, however, a marginally significant difference between Control and Gettier (Fisher's exact test p=.07) (Starmans and Friedman, 2012: 275). The binomial tests show that participants are not answering at random, but in response to real features of the vignette. The large majorities attributing knowledge in both Control and Gettier suggest that both are typical cases of knowledge according to the folk concept. The comparison of Control and Gettier suggests that for some participants, at least, there is some difference between the vignettes. Starmans and Friedman do not linger on this fact, and they in fact go to some lengths to minimize the difference. This is disappointing. Although it is clearly an *exciting* result, given the philosophical consensus, to report that lay people do not share the Gettier intuition, by minimizing the difference they *did* find, Starmans and Friedman seem to be leaving unexplored what are, perhaps, interesting complications for the theory of knowledge attribution. It surely fun

concept of knowledge licenses knowledge attribution to agents in both uncontroversial cases and in Gettier cases, but not to agents with false beliefs.

These results help solve the control problem facing WNS. Starmans and Friedman's study also begins to address the generality problem, as they test two further variations of this design, one in which the thief takes a pen, and then accidently drops it; in another an employee takes money out of an open cash register, and then accidently drops money into it. The results from both these tests are consistent with the ones described above.[61] Starmans and Friedman conclude that,

> Probably the most striking finding from the current experiments is that laypeople readily attribute knowledge in Gettier cases. Whether this seems surprising will probably depend on one's own intuitions about whether knowledge is possessed in Gettier cases. But this finding suggests that laypeople have different intuitions from philosophers, who apparently hold with near unanimity that Gettiered agents are not knowledgeable.[62]

They consider two related possible explanations of the differences between philosophers and laypersons, drawn from the literature on experimental philosophy, without explicitly endorsing any one in particular. The first explanation they consider is that the difference is like the one between a novice and an expert.[63] According to this line of thinking, philosophers and lay people have the same concept, but differ in their ability to apply it to cases. I'll table this explanation for the time being because it is one of the major ways in which traditional philosophers have attempted to resist the variation argument. A second possible explanation is that laypeople and philosophers do not have the same intuitions because they do not have the same concept or conception of knowledge. According to this account, there is no difference in ability, just a bare difference in conceptual scheme.

Why might this be? One possibility is the selection effect in graduate education suggested by WNS. Another, also suggested by WNS, is that philosophical training actually changes philosophers' concept of knowledge. Perhaps the invocation to lead an examined life and intensive training leads to a more stringent concept of knowledge and more

---

to say "philosophers were wrong" but I think it's more interesting to try to determine i) what factors turn Gettier intuitions on, what factors turn them off, and whether there are ii) individual differences in sensitivity to them

[61] *ibid,* 276 - 278.

[62] *ibid*, 280.

[63] See E. Sosa (2007), and Weinberg, Gonnerman, Buckner, and Alexander (2010), for the first discussions of this possibility.

demanding standards of epistemic assessment. If philosophical training really does change a person's concept of knowledge, this could easily lead to different intuitions about philosophical thought experiments.[64] Another possibility, that I'll explore in Chapter 5 in connection with skeptical cases, is that rather than changing a philosophers' concept, training changes the broader evaluative strategies they employ, and their sensitivity to the cues that engage the strategy.

4.6 The Reformed General Variation Argument

If the explanation of the difference between laypeople and philosophers' intuitions about Gettier cases discovered by Starmans and Friedman is that the groups have different concepts of knowledge, this would underwrite a strong version of premise 3, and pose a serious challenge to traditional methodology.[65] Recall that WNS's concern was that the method of cases relied on "local" intuitions that varied between groups. The goal of their

[64] See S. Stich and W. Buckwalter (forthcoming).

[65] Starmans and Friedman's results are "hot," so to speak. A number of papers are presently in press at top journals contesting these results. Jennifer Nagel, Valerie San Juan, and Raymond A. Mar (forthcoming a) have run a series of experiments showing that "[a]lthough participants rated true beliefs in Gettier and Skeptical Pressure cases as being justified, they were significantly less likely to attribute knowledge for these cases than for matched True Belief cases. This pattern of response was consistent across different variations of Gettier cases and did not vary by ethnicity or gender… These findings therefore suggest that across demographic groups, laypeople share similar epistemic concepts with philosophers, recognizing a difference between knowledge and justified true belief" (pg. 1 mnscpt). Most of Nagel et al.'s experimental material was new, but they did rerun the original vignettes from WNS, using Canadian undergraduates with similar ethic backgrounds and Nagel et al. failed to find a significant effect (pg. 7 mnscpt). An important caveat is that null results, especially those arrived at with the use of inferential statistics, are generally considered inconclusive. This is a serious concern here as Nagel et al.'s sample sizes for the South Asian and East Asian comparison classes were relatively small, and it's far from clear that their *N*s gave them enough power to detect a possible difference. The same point about accepting the null applies to Nagel et al.'s more general conclusion. Failure to detect a difference does not imply that there is no difference, otherwise Newtonian experiments would be sufficient to contradict relativistic mechanics. Their data does nonetheless provide some evidence that laypeople do, or can be prompted, to have the Gettier intuition. The very same point about accepting the null *also* applies to Starmans and Friedman. They gain some generality for their claim by trying a number of vignettes, but they still cannot infer that no difference in *these* tests really shows that laypeople are different from philosophers for the reasons just mentioned, and because they must rely on philosophers' personal assessments of Starmans and Friedman's experimental stimuli as authentic Gettier Cases. Assessing the relative merits of Starmans and Friedman versus Nagel et al. must involve a more thorough weighing of their respective methodological merits; this is a discussion that would be taxing in an already overly long footnote. It deserves to be noted that in an even more recent exchange Starmans and Friedman (forthcoming a) have argued that Nagel et al.'s (forthcoming a) results use non-standard Gettier cases, and that their presentations are biased to produce knowledge denials. Nagel et al. (forthcoming b) have responded that their vignettes are in fact more central to the philosophical literature, and they provide new experimental evidence that their methods are not biased toward knowledge denial. Assuming, as is proper, that these experimental results are sound, what we may have in the offing is an *experimental* rehashing of the Gettier analysis debates of the 70's and 80's, the end conclusion of which may be that there is no single *canonical* "gettier" problem, and thus no single possible analysis, and that the proliferation of cases taps into a complex web of features, all relevant to knowledge attribution.

experiments was to establish this. The group divisions they focused on were relatively large ones, like those between the Eastern and Western hemispheres, and those between rich and poor. If, in fact, the relevant division between intuition groups lies at a much more local level like that between professional philosophers[66] and laypeople, then Starmans and Friedman's results do support a strong version of the variation argument upon which what it is to *know* is different for different groups. If what it is to know is really different for different groups, then the difference in concept will lead to different norms and virtues of inquiry. There is nothing inherently incoherent about this position, but it is for many quite objectionable for the following familiar reason. If norms of responsible inquiry really do differ from group to group, then the conclusions of different groups across many domains of human interest may vary as well. And then there would be no questioning the conclusions of other groups. If Starmans and Friedman are right about the location of the division in epistemic intuitions, this underwrites a dramatic variation argument. The intuitions of philosophers in the analytic tradition are objectionable, on this version, not because they are the voice of rich Westerners, and the rich and Westerners tend to be different from the poor and non-Western, but because philosophers are just different from everyone else! It's a much more immediate difference, and philosopher's intuitions speak for a group that is a very much smaller minority. Thus, to reformulate the general variation argument: If intuitions are evidence for norms, and intuitions between philosophers and laypeople differ, then philosophers and laypeople have different norms. The dilemma a reformed version of WNS's challenge poses is between embracing relativism and rejecting the method of cases.[67]

---

[66] This argument does rely on the assumption that *philosophers* share the Gettier intuition, but see T. Williamson, 2005, for the claim that it is widely and more or less immediately shared among the group. In support of this, in a survey of professional philosophers, the results of which are reported in "What philosophers Believe" (forthcoming), David Bourget and David Chalmers asked more than 900 philosophers 30 questions about their views about questions drawn primarily from the recent literature in Mind and Epistemology. They did not ask about Gettier cases, presumably because they believed a consensus was already clear.

[67] This worry is akin to one that has been recently discussed across the social sciences. In an influential and widely discussed paper, Joe Henrich, Steve Heine, and Ara Norenzayan (2010) argue that at the heart of the social sciences is a quite dubious practice. These fields tend to issue broad claims about humanity, and human psychology, but do so based on studies with experimental samples drawn from a population of Western, Educated, Industrialized, Rich, and Democratic (WEIRD) 18-22 year old college students with above average IQ, and who mostly attend elite research universities. While not the least representative sample of humanity, Princeton undergrads come pretty close. It's plausible to suppose, according to Henrich et al., that if the social sciences used samples with more cultural, educational, socio-economic, and political variation that many of the signal results would be quite different. There's a quite charitable interpretation of the practice, of course. Experimenters begin by assuming that there are fundamentally no differences between the Princeton undergrad

4.7 Experimental Philosophy and the Skeptical Problem

Although the Gettier problem is sometimes thought to be related to the skeptical problem, most philosophers treat them separately.[68] So, why think variation in Gettier intuitions is anything more than a local phenomenon? Perhaps *Gettier* intuitions vary across groups, but if Gettier cases and the skeptical problem are distinct, why think that the intuitions which drive skepticism might be different among non-philosophers? The prima facie worry is that we do not know why philosophers' intuitions differ from folk intuitions in Gettier cases. Not knowing why, we don't know where or when they might diverge. While many philosophers share the Gettier intuition, they've offered many incompatible analysis of knowledge in response. This shows that the structures which drive intuitive responses are opaque; we cannot recover them merely be reflection, so we must instead try to reconstruct them. So, it's a substantive question why they differ. Of course it's not the case that this disagreement is reason to mistrust all your intuitions -- no one need doubt their intuitions about simple addition -- but given that the topic is the same, *knowledge*, we have reason to worry.

While the Gettier intuition is by far the most widely studied in experimental epistemology, there is some empirical work that bears on AI, and supports the idea that it deserves a more systematic empirical treatment. There are three areas where recent studies suggest that intuitions related to skeptical arguments might vary, one concerning Dretske's Zebra cases, another concerning direct tests of radical skepticism, and recent work on salience effects and linguistic contextualism. I'll take these up in turn.

We've already seen that one study by WNS showed that intuitions about Dretske's Zebra case vary across socio-economic groups. In a separate study, Shaun Nichols, Stephen

---

and the Napalese herder because it would be chauvinistic to do otherwise. Even as it's clear to see that the presumption of uniformity begins with an interest in avoiding the kind of objectionable cultural self-importance that marred western thinking for so long, it nevertheless assumes the very kind of result that is often at issue in such experimental studies. Often we want to know if some feature, some tendency, some way of thinking is a universal or a product of culture. We want to know if it is innate or learned; or to borrow a metaphor from Gary Marcus, to what extent culture can revise nature's first draft. But anti-chauvinism assumes the answer to this question. Henrich et al.'s critique of the social science practice differs from WNS's critique of philosophical practice in import respects. The social sciences are descriptive. They aim to uncover facts and construct accurate theories of human behavior from them. They do not issue normative conclusions, and consequently the style of dilemma WNS construct does not emerge from their argument. Their conclusion is simple: the social sciences should expand their experimental base.

[68] Stephen Hetherington (2012) has argued that Gettier cases are relevant alternatives error cases, with a bit of good luck added at the end. I think there are good reasons to believe he's right, but the view is nevertheless not a standard one.

Stich, and Jonathan Weinberg (NSW) found cross-cultural variation in response to a similar Zebra case.[69] They found that approximately 70% of Westerners shared the philosopher's intuition that the agent in Dretske's vignette "only believes" that the animal is a zebra. By contrast, they found that among participants with a South Asian background, only 50% thought the agent "only believes." The difference is statistically significant. If this relevant alternatives intuition plays a role in the intuitive grounding of premise one in AI, then this study offers some evidence for variation in the intuitions that drive the skeptical problem.

The same methodological worries I aired in section 4.4 apply here as well, so the case deserves further experimental scrutiny to be sure. But supposing that those worries are not serious, one could wonder if this only shows that the appeal of skepticism is culturally local, and that the skeptical problem is *a problem* not for philosophers, but for Westerners. For it seems, at least according to this result, that non-philosophers within the western cultural tradition do have the relevant alternative intuition, so we don't yet have reason to mistrust the traditional description of the skeptical problem.

NSW focus on this question more directly in another study. They presented two groups of undergraduates, one with a background in philosophy and another without, with a brain-in-a-vat scenario. The study used their familiar format, it presented both groups with a scenario, and then asked if the agent in the scenario "really knew" that he is not a brain in a vat, or does he "only believe" it. The scenario they use is as follows,

> George and Omar are roommates, and enjoy having late-night 'philosophical' discussions. One such night Omar argues, "At some point in time, by, like, the year 2300, the medical and computer sciences will be able to simulate the real world very convincingly. They will be able to grow a brain without a body, and hook it up to a supercomputer in just the right way so that the brain has experiences exactly as if it were a real person walking around in a real world, talking to other people, and so on. And so the brain would believe it was a real person walking around in a real world, etc., except that it would be wrong – it's just stuck in a virtual world, with no actual legs to walk and with no other actual people to talk to. And here's the thing: how could you ever tell that it isn't really the year 2300 now, and that you're not really a virtual-reality brain? If you were a virtual-reality brain, after all, everything would look and feel exactly the same to you as it does now!"
>
> George thinks for a minute, and then replies: "But, look, here are my legs". He points down to his legs. "If I were a virtual-reality brain, I wouldn't have any legs really – I'd only really be just a disembodied brain. But I know I

[69] S. Nichols, S. Stich, and J. Weinberg, 2003.

> have legs – just look at them! – so I must be a real person, and not a virtual-reality brain, because only real people have real legs. So I'll continue to believe that I'm not a virtual-reality brain."
>
> George and Omar are actually real humans in the actual real world today, and so neither of them are virtual-reality brains, which means that George's belief is true. But does George know that he is not a virtual-reality brain, or does he only believe it?
>
> REALLY KNOWS ONLY BELIEVES

They found that while approximately 80% of students with philosophical training thought that the agent "only believes" they're not a BIV, only 45% of those without philosophical training shared that intuition.

This study provides some evidence for the idea that philosophers' intuitive response to the skeptical problem differs from non-philosophers'. It is only the beginning of an empirical description of the skeptical problem though. There are a number of ways the test could be refined to reveal exactly what drives the difference between the two groups. First, if this prompt is tapping into the same intuitions that AI does, this study won't tell us where the difference lies. Is it the error intuition? The closure intuition? The intuition of conflict? All three? It's possible the difference can be explained by those in the philosophical group lacking the ordinary intuition. A further difficulty is that this prompt pairs a narrative version of an error argument for skepticism with a version of a Moorean argument for dogmatism. It's quite possible that both groups were equally moved by the skeptical argument, but the non-philosophy group found the Moorean argument more compelling than the philosophy group.

Another difficulty with interpreting these results is assuming that the "philosophical training" of one of the groups, which for NSW amounts to having had three previous classes in the subject, is really comparable to that of professional philosophers. Because NSW use a stock scenario it's quite reasonable to think that the responses of the trained group were accommodation rather than genuine reports: they may have been better positioned to guess what the experimenters "wanted" them to say, and responded accordingly, irrespective of their genuine intuitions.

A better empirical approach would be more systematic: it would look at each of the intuitions driving the philosophical description of the skeptical problem in isolation. Doing

so would allow a much more nuanced evaluation of the intuitive basis of the skeptical problem, indicating where and even how individuals differ. This is a task I'll take up in Chapters 3 and 4. In sum, while I have some reservations about this study, and how much light it can really shed on the details of the skeptical problem, those reservations don't offer much comfort to armchair descriptions of the skeptical problem: this study suggests that even among Westerners, there is variation, of some kind, in response to skeptical arguments.

A final reason to worry about armchair diagnosis comes from recent empirical investigations of linguistic contextualism. Linguistic epistemic contextualism (herein just contextualism) claims that "know" is a context-sensitive term and, consequently, the truth-conditions for sentences that embed "know" can vary according to context. A number of authors have recently applied this idea to providing a diagnosis of radical skepticism. I'll have much to say about the view in subsequent chapters, but here's the gist of the view. According to contextualists if "know" is context sensitive, then the sentence "I know I have hands" can be true when uttered in a context with low epistemic standards, while the same sentence, in the mouth of the same speaker, can be false in when uttered in a context with high epistemic standards. What changes contextual standards? Many things according to contextualists, but one important conjecture is that it can be accomplished by raising possibilities of error that haven't, or can't, be ruled out. This, at least, is the contextualist's explanation of the error intuition. According to the theory, the skeptic raises the standards for knowledge by introducing an error possibility, and because we cannot rule out that we're a brain in a vat, or deceived by a demon, we have the intuition that we don't know that we're not.

Contextualism is an empirical thesis about the semantics of "know" that has been applied to the skeptical problem. But why think that "know," a verb, is context-sensitive, when it has so little in common with more familiar examples, including indexicals like "I," demonstratives like "this," or scalar adjectives like "flat?" The basic source of evidence for contextualists are intuitions about cases that seem to demonstrate that the conditions of application for "know" can shift depending on whether certain alternatives have been made salient. To take an example, contextualists like Keith DeRose point to pairs of cases like the following, in Table 3.

Low Standards	My wife and I are driving home on a Friday afternoon. We plan to stop at the bank on the way home to deposit our paychecks. But as we drive past the bank, we notice that the lines inside are very long, as they often are on Friday afternoons. Although we generally like to deposit our paychecks as soon as possible, it is not especially important in this case that they be deposited right away, so I suggest that we drive straight home and deposit our paychecks on Saturday morning. My wife says, 'Maybe the bank won't be open tomorrow. Lots of banks are closed on Saturdays.' I reply, '**No, I know it'll be open**. I was just there two weeks ago on Saturday. It's open till noon.' (emphasis added)
High Standards	My wife and I drive past the bank on a Friday afternoon, as in case A, and notice the long lines. I again suggest that we deposit our paychecks on Saturday morning, explaining that I was at the bank on Saturday morning only two weeks ago and discovered that it was open until noon. But in this case we have just written a very large and very important check. If our paychecks are not deposited into our checking account before Monday morning, the important check we wrote will bounce, leaving us in a *very* bad situation. And, of course, the bank is not open on Sunday. My wife reminds me of these facts. She then says, 'Banks do change their hours. Do you know the bank will be open tomorrow?' Remaining as confident as I was before that the bank will be open then, still, I reply, '**Well, no, I don't know.** I'd better go in and make sure.' (emphasis added)

Table. 3.

DeRose invites his readers to render intuitive judgments about whether the agents in the vignettes have spoken truly, and thus whether or they possess knowledge.[70] DeRose's intuitive verdict, which is widely shared among philosophers, is that in both vignettes the agent speaks truly. If so, it follows that the agent in the first vignette know the bank will be open, while the agent in the second vignette does not, even though the agents have the same information about the bank's hours.[71] This shift in intuitions based on whether or not an error is salient is sometimes called a salience effect.[72]

Linguistic contextualism is one of the leading, if not the leading, contemporary approach to diagnosing the skeptical problem. If salience effects play a role in the appeal of

[70] K. DeRose, 1992: 913.

[71] This pattern of intuitive salience effects is widely accepted in contemporary epistemology, including by anti-skeptical invariantists (P. Rysiew, 2001; J. Hawthorne, 2004; T. Williamson, 2005; A. Hazlet, 2009; and J. Nagel, 2010); skeptical invariantists (P. Unger, 1971); subject-sensitive invariantists (J. Stanley, 2005); and of course linguistic contextualists who are in focus here (G. Stine, 1976; D. Lewis, 1979 & 1995; S. Cohen, 1999; J. Shaffer, 2004).

[72] DeRose also changes the stakes of error in these cases. There is a lively contemporary debate over issues of pragmatic encroachment, and I discuss them in more detail in Chapter four.

premise one in AI, then the contextualist's linguistic conjecture about "know" plays an essential role in the theory, and of their description of the skeptical problem. Recently, however, a number of attempts to empirically study salience effects have shown that non-philosophers do not seem to have the salience intuition: mentioning an unrealized possibility of error does not undermine their willingness to ascribe knowledge to an agent. If philosophical consensus does differ here, then there is a large reason indeed to worry about armchair diagnosis.

Two recent studies of salience effects, one by Wesley Buckwalter, and another led by Josh May, failed to find any evidence that mentioning an unrealized possibility of error changed non-philosophers' intuitive knowledge ascriptions.[73] Both studies used prompts very similar to those used by DeRose, and sampled large populations. I'll have much more to say about these studies, and salience effects more generally in Chapter 4, but the for the moment it is enough to notice that these studies put further pressure on the armchair description of the skeptical problem, and one prominent diagnosis of it. The error intuition is central to the skeptical problem because the skeptic is not defending the idea that we are deceived—that would amount to delusion not doubt—only that it is possible that we are deceived. If our ordinary concept of knowledge does not require ruling out salient counter hypotheses, then its open to wonder how the skeptical threat is intuitive at all; it is certainly not plausible. If non-philosophers do not share the error intuition, this is a serious difference between philosophical and non-philosophical epistemic practice.

To sum up, I think there is a real reason to worry that the epistemological intuitions of philosophers and non-philosophers differ, at least in some areas. I think there is real reason to worry that philosophers' and non-philosophers' intuitions about the skeptical problem differ. Early restrictionists like Weinberg and Stich concluded that this was reason to doubt that appeals to intuition could serve to ground normative theories, but my concern is different. Our intuitions are evidence for nothing more than our ordinary epistemic practices. Uncovering the factors that influence them helps us describe the psychological mechanisms that subserve our intuitions. Studying the intuitions that give rise to the skeptical problem is a descriptive task; its goal is to illuminate the conditions under which we feel the threat of the skeptic, and the conditions under which we don't. If the conditions

[73] W. Buckwalter, 2010; May et al., 2010.

under which philosophers feel the threat of skepticism differ from those under which non-philosophers feel the threat, then both belong to the descriptive project.

But here is where description becomes essential. Description is prior to diagnosis: responding to the skeptic's threat depends upon understanding it. Different descriptions will lead to different diagnoses, and different kinds of options in response to the skeptic. Any successful response to the skeptic depends upon getting the description right. I'll have much more to say about this issue in the chapters to follow, but there is a clear connection between description and normative response to the skeptic. Linguistic contextualism is concessive, and it is concessive because its diagnosis of the appeal of salience effects points to their foundation in ordinary epistemic practice. Some anti-skeptical invariantists trace the appeal of salience effects to cognitive biases, and they argue we need not accept them because of this. In this way description can drive one's normative stance.

5. Responses to the Variation Argument

Up until this point I think we should be tentative in our conclusions. We don't have decisive evidence one way or another that there is any kind of between group variation in these intuitions. The extant studies are too few, they are not systematic, and they fall far short of demonstrating the generality of their conclusions. But they are reason to worry that intuitions differ, and reason to conduct further more systematic empirical studies. Nevertheless, perhaps there are other ways to deflect this empirical challenge. In this section I'll consider three kinds of response that have been offered in defense of more traditional appeals to intuition, and weigh their merits as justifications of the armchair description of the skeptical problem.

Perhaps the most popular response to the challenge from the variation argument is the *expertise defense.* The basic idea is to accept that the intuitions of philosophers and non-philosophers may differ, but reject that this is a cause for worry. Kirk Ludwig develops the defense in the following way,

> We should not expect antecedently that untrained subjects should be in an especially good position to give judgments in response to scenarios involving difficult questions…What is called for is the development of a discipline in which general expertise in the conduct of thought experiments is inculcated and in which expertise in different fields of conceptual inquiry is developed and refined. There is such a discipline. It is called philosophy. Philosophers are best suited by training and expertise to conduct thought experiments in

> their areas of expertise and to sort out the methodological and conceptual issues that arise in trying to get clear about the complex structure of concepts with which we confront the world. A lot of the problems we confront are very difficult, and so it is not surprising that in many areas there is lively debate, but it would be a mistake to think that the way to resolve such debates is to return to questioning untutored subjects in just the places where there is evidence of the most difficulty in coming to a clear view.[74]

Steven Hales offers a critique along similar lines. He argues that intuitions are responsive to training and education, and consequently physicists' intuitions about physics are better than their students', and mathematicians' intuitions about mathematics are better than those of the average joe. He concludes that "philosopher's intuitions are much more reliable than either those of inexperienced students, or the 'folk'."[75]

So, if there are expert intuitions, why do some people have them? One explanation is that philosophers have better concepts. This might be true when it comes to logical concepts: philosophers just have clearer, more consistent concepts of things like validity and implication, and it's why we should trust their assessments of whether an argument is valid, or whether one statement implies another, more than folk intuitions about the same. Or, perhaps philosophers and non-philosophers have the same concept, but philosophers have more precise conceptions, and this aids their judgments. But do philosophers have a better concept (or conception) of knowledge? If by better we mean different, than the expertise defense offers little comfort to diagnostic approaches to the skeptical problem. They are premised on the idea that skeptical arguments appeal to something in our ordinary concept of knowledge - recall Bernard Williams here: "our ordinary concept of knowledge invites skepticism." Embracing this version of the expertise defense precludes them from offering an account of the impact of skepticism on ordinary knowledge. Michael Williams describes this kind of high standards *philosophical* skepticism as uninteresting because it does not tell us why we should adopt the hyperbolic standard of the philosopher.

This is, in fact, a serious worry for the armchair analysis of the skeptical problem, one that has been lurking in the shadows of my discussions of experimental results up until this point. It is not implausible to think that philosophical training results in philosophers developing quite different concepts. If epistemologists' training results in their having a

[74] K. Ludwig, 2007: 150 - 1.
[75] S. Hales, 2006: 171.

different, perhaps more stringent, concept of knowledge than one finds ordinarily, it could result in a warped description of the skeptical problem. This itself is a plausible diagnosis of the difference in philosophers' response to the skeptical problem and it is consistent with anecdotal evidence about how students first react upon encountering Descartes' demon.

Perhaps another route is to suggest that philosophers don't have different concepts, just a better understanding of our ordinary concept. Joshua Knobe and Shaun Nichols consider a reply along these lines, offered in response to experimental philosophy in general, but they reject the idea that it renders experimental investigations irrelevant. If philosopher's understanding of concepts like knowledge differ from ordinary people's, it's an interesting project to understand why and how, and the best way to do that is experimentally. They go on to argue that, "even if we discover important difference between philosophers and the folk, it would hardly follow that data from the folk are irrelevant. Rather, the whole pattern of data might tell us something important about the ultimate source of the philosophical problem."[76]

Both kinds of expertise defense seem to recommend more experimental work, not less. In some cases, like contextualism, the diagnosis of the skeptical problem is premised on an explicitly empirical claim, in the present case about the semantics of "know." Keith DeRose is explicit about this, saying "wise contextualists, however, as we have seen, seek to support their contextualism by appeals to what transpires in ordinary, non-philosophical uses of 'know(s)', the most important examples of which are cases that involve no dispute among the parties whatsoever, before *applying* their contextualism to the problem of philosophical skepticism."[77] If "know" is not context-sensitive in ordinary usage, then the contextualist diagnosis of the skeptical problem should be rejected. The expertise defense is not an option, and is in fact self-undermining.

Another kind of response is the *no intuitions* defense. According to this view all the talk of intuitions, understood as immediate responses, be it as evidence or otherwise, is really beside the point because the philosophical method doesn't really employ intuitions. Herman Cappelen offers a version of this defense in his recent book *Philosophy without Intuitions.* He argues that philosophers use the word "intuition" loosely, and rarely if ever use intuition as evidence. He is charmingly unequivocal in stating the position, "on no sensible construal of

[76] J. Knobe & S. Nichols, 2008: 9.
[77] K. DeRose, 2009: 57.

'intuition', 'rely on', 'philosophy', 'evidence', and 'philosopher' is it true that philosophers in general rely on intuitions as evidence when they do philosophy."[78] But here his central idea is that intuitions serve as evidence for particular claims. So, when Gettier describes the case of Smith and claims that it is "clear" that Smith doesn't know, this functions according to some as evidence for the claim that knowledge is not justified true belief. This construal of the use of intuition in Gettier cases and other epistemological thought experiments, like Keith Lehrer's true-temp (to be discussed in a moment) or Alvin Goldman's fake-barn cases, is what Cappelen objects to. In particular, he offers a deflationary account of intuition-talk, contending that examples like these and our responses to them are best understood as establishing what's in the pre-theoretic common ground, and that non-intuitive argument is required to explain why our responses to these cases should be part of the common ground.[79]

While I have some concerns with Cappelen's over-arching thesis, I don't think it obviates the need for empirical work in describing the skeptical problem. As we saw in the preface, many of the main participants in the discussion of the skeptical problem in the last quarter century explicitly describe it as an intuitive problem. But more to the point, the intuitions to which they appeal don't function as evidence for or against a particular theory, but as part of a description of a problem. So, in Cappelen's terms they serve to establish what's in the common ground. What remains, then, to set up the skeptical problem is to point out that those common ground commitments are inconsistent. It is the project of further theorizing and no doubt argument, which may or may not rely on intuition, to determine what part of the common ground to give up. But the descriptive question that's at the heart of the diagnostic project is about what is, and what is not, part of the common ground when it comes to our ordinary epistemic practice and concept of knowledge. The worry that I'm outlining is that these are different. Establishing whether or not they're different has nothing to do with whether intuitions are evidence, but with gathering evidence about intuitions.

A third kind of response to the deflationary results of experimental philosophy is to distinguish between experimental participants' *responses* and their actual *intuitions*. According to the view, which I'll call the Intuition-Response Defense (IRD), experimentalists assume

[78] H. Cappelen, 2012: 3.
[79] H. Cappelen, 2012: 167 - 72.

without adequate evidence that the results they gather reflect individuals' actual *intuitions* about a target, like knowledge or reference, and they do so without ruling out that those responses are guesses, explicit theorizing, or any manner of non-intuitive response. This is because experimental studies, whether conducted online or on paper, typically present different groups stimulus materials that are then followed by a question, or some other kind of scale, to illicit a response. Experimental studies don't observe intuitions; they infer them. IRD concludes, then, that experimental philosophy does not show that intuitions are unstable, or that folk intuitions differ from philosophical intuitions, just because it's not clear that experimental philosophy is comparing intuitions to intuitions; it may be comparing folk responses (guesses, naïve theorizing, memories, and the whole dog's breakfast of mental states if you like) to philosophical intuitions. And from that comparison nothing untoward follows about philosophical intuition. Applying this as a defense of the philosophical description of the skeptical problem might go something like the following, "sure, folk responses to the skeptical problem differ from philosophers' because they're looking at cases and arguments that are unfamiliar and hard to understand, so it's no wonder if they're responses are different. They're just guessing rather than applying their concept of knowledge to these cases."

The worry is worth considering. Notice that this is something of a specialized variant of the expertise objection. It doesn't posit that philosophers have better concepts, or even a better understanding of their concepts, than ordinary folk, just that philosophers can recognize their own mental states better. Because of this they know they're appealing to intuition rather than just guessing. By contrast, empirical studies don't tell us about the mental states of participants, just what they mark down on a paper, or what button they press on a computer.

John Bengson frames a clear version of the IRD in a recent paper. He argues as follows,

> experimental attacks on intuitions…neglect a considerable gap between the answers elicited by the relevant empirical studies and the intuitions about which naysayers naysay. It cannot innocently be assumed that subjects' answers expressed how things struck them—what intuitions they had, if any.[80]

[80] J. Bengson, 2013: 496

Bengson goes on to argue that we have no reason to think that guessing and naïve theorizing are uncommon, so we have no reason to think, and perhaps reason to reject, the idea that survey answers typically reflect participants' intuitions.[81]

Bengson's worry is well taken, but it threatens to prove too much. Experimental philosophy simply takes over the methods of the social sciences, which investigate all manner of phenomena, including reasoning, categorization, the structure of concepts, among much else. It would be easy to transpose his worry to these kinds of investigations as well: studies of reasoning might not get at participants' actual competence, but rather illicit guesses or naïve theories of reason. One might perhaps worry about the whole of the social sciences, but it would be better if there were reason to think that the applications of these methods of philosophy were particularly prone to eliciting guesswork.

But perhaps there is a way of focusing the challenge. A concern that is often raised is that the kinds of scenarios used by philosophers, and then deployed in surveys, are often underdescribed and this could lead to participants responding at random. The thought goes something like the following. Professional philosophers are able to fill in the details because they have the right kind of training, and they're familiar with the larger discussion. But non-philosophers don't have this background and can't fill in the details, so they're left wandering in the dark, grasping and guessing at random. Simon Cullen argues that this in fact leaves participants vulnerable to misinterpreting the scenarios they're being asked to consider.

> [F]ollowing in the tradition of contemporary thought-experiments, experimental philosophers have invented some stunningly bizarre survey vignettes…Convoluted thought-experiments are ripe with opportunities for confusing subjects, breaching conversational norms, and inadvertently conveying information which, in their effort to provide intelligent responses, subjects mistakenly interpret meaningfully.[82]

The work of Norman Schwarz has shown that fine contextual clues can lead participants to systematically misinterpret the intent of social science experiments.[83] Thus, introducing a bizarre or far-fetched scenario, such as demon deception in the study of skepticism, might

[81] Bengson is a strong rationalist, and his view is that intuitions are a distinctive kind of mental state, and in particular that they are warrant-conferring, non-derivative intellectual seemings. But nothing in IRD really depends on a commitment as particular as this. All that matters for his part is that there are guesses, judgments, and some third category, intuition, which is a non-inferential non-guesswork assessment.

[82] S. Cullen, 2010: 7 (online first mnscpt.).

[83] N. Schwarz, 1995 & 1996.

only exacerbate the problem. If Cullen is right, it might offer a way of focusing Bengson's worry without proving too much.

So, if the expertise defense and the no intuitions defense are inimical to the diagnostic approach to the skeptical problem, then IRD might be the best way to resist the kind of empirical approach I advocate here. Should we worry that non-philosophers' intuitions about the skeptical problem are somehow guesses? There are simple statistical tests to determine whether responses are merely random, so the most serious worry is that they're somehow systematic guesses, or systematic naïve theorizing. I think there are two related ways to address these concerns, and I'll take them up in turn.

Are experimental scenarios too thinly described? Although it depends on the case, I think this is unlikely. Chinua Achebe has observed that wherever you go in the world, telling stories is a human universal. It's not that we all like to tell stories, but that telling stories has its routes in sharing social information. Sharing social information about members of our social group served many important evolutionary functions, as navigating the cooperative landscape of communal life was one of the primary adaptive problems that faced our ancestors. Knowing what others know, knowing what others believe, knowing when others have mistaken beliefs were all important accomplishments because they would allow an individual to predict and explain the behaviors of their conspecifics. Robin Dunbar has conducted a wide-ranging evolutionary and anthropological study of human communication, and has hypothesized that the *primary* evolutionary function of communication was to share social information. Indeed, in a number of studies of contemporary linguistic behavior, he's found that upwards of 75% of all human communication is *gossip*, that is communication about the actions and mental states of others in our social group.[84] The overall picture is something like this: primates live in groups to protect themselves from predation and secure the advantages of cooperation; social living also involves many adaptive problems, both of in-group competition and cooperative free-riding; so, living in social groups depends on the ability to "understand and exploit the mental states of others."[85] In other words, it requires a theory of mind, a capacity that many believe is innate. Furthermore, all you need to do is remember the last conversation you had with a group of more than a few people, and I think it will be clear how fragmented, interrupted, abrupt, and circuitous gossip can be: before you

[84] R. Dunbar, 2004.
[85] R. Dunbar, 2004: 103.

even finish your story about what X did, it reminds me of something Y did that I just have to say quickly. Gossip is just a form of story-telling, and understanding the mental lives of agents in stories is something that we are very, very good at. Our survival depended on it.

Focusing on epistemology now, the kinds of scenarios that are deployed in experimental studies are just short stories. They ask study participants to interpret what's going on, and in the case of epistemology to track the mental states of the agents in the stories, what they know, and what they merely believe. This is something we do automatically. This automatic assessment is just intuitive cognition: it is automatic, effortless, and thus reflects our ordinary epistemic competences. Sometimes when we read fiction, especially mystery novels and the like, we do have to explicitly ask ourselves what the characters knew, and when. But that's the exception, not the rule. For the most part tracking the mental states of agents in stories its something we do automatically, and without effort or conscious direction. Experimental epistemology only asks people to exercise these same capacities, and to sometimes report the automatic assessments they're making. Far from thinking that survey procedures are likely to draw forth guesswork and naïve theorizing, at least when it comes to epistemology and the exploration of knowledge and belief, we have very good positive reasons to think that survey responses are very likely to be accurate reports of our intuitive assessments, and consequently that they reflect our underlying ordinary epistemic competence. And this is one place where more and better empirical work has its place - just as Joshua Knobe and Shaun Nichols called for in the passage I cited above. The task is to study particular questions in depth, and repeatedly, to see what kinds of intuitive responses are stable, and what kinds of factors can change them. If we find stable kinds of responses, and we find stable factors that can affect them, then given our prima facie competence at understanding mental states in social story telling, the best explanation is that those regularities reflect our underlying competences, and not guess work.

But what about the peculiar thought experiments of philosophers? Perhaps our intuitions will reflect our ordinary epistemic competences when it comes to normal cases, but fake-barns, true-temp thermometers, and above all demon deception might just be too far afield. These kinds of far-fetched tales might invite reflective intervention that screens off intuitive monitoring. I think this is a genuine worry, and not one to reject out of hand. But the proper response to a worry like this is not to give up on the descriptive project; it's

more and better empirical investigation. It's an empirical hypothesis that philosophers' curious researches darken our intellect. But it's one that can be controlled for, and investigated systematically with careful studies. And this is just the challenge I want to take up in the next two chapters. One way to sum up my response to the bizarreness objection is this: if the bizarreness objection seems like a serious reason to think that philosophers and non-philosophers differ, the more urgent the call to do empirical work to test the hypothesis. I think philosophical training is a good thing - but perhaps it is not always the best kind of training for appreciating our ordinary capacities. Assume for a moment that there is a difference between philosophers' and non-philosophers' intuitions about the package of intuitions behind the skeptical problem. It's a further question which intuitions we should prefer. Describing the differences in intuitions is just a first step, and is not yet an evaluation of them.

6. Conclusion

In Chapter 1 I presented the skeptical argument from ignorance, and the consensus view that it is an intuitive paradox that reveals an inconsistency in our ordinary practices of epistemic assessment. I then argued that recent work in cognitive psychology raises the worry that our intuitive assessment do not always conform to our best normative theories, and that they are not always consistent. I argue that these results are good reason to take an empirical look at the intuitions driving the skeptical problem. In this chapter I've introduced experimental philosophy, and I've argued that recent work on folk theories of knowledge, and on thought experiments related to the skeptical problem suggest that philosophers and non-philosophers may have different intuitions about the skeptical problem. If they do, then the consensus description of the skeptical problem we began with may be inaccurate, incomplete, or both. I then considered a number of objections to experimental philosophy, and the possibility that non-philosophers' intuitions are not relevant to the skeptical problem. I argued that they are, and that the kinds of reasons typically given for ignoring the folk result in a kind of high-standards skepticism that most philosophers do not find threatening. The most threatening versions of the skeptical problem begin in our ordinary conception of knowledge, and as such, folk intuitions belong to its description. But since a compelling response the problem of skepticism, especially one that will explain how we fell into its trap, requires an accurate description of the skeptical problem, it follows that we

must look at the skeptical problem in a systematic way that takes seriously the intuitions of non-philosophers. And so it's toward the descriptive project we now turn.

Chapter 3

Contextualism & the Intuition of Conflict

1. Introduction

The focus of this chapter is linguistic epistemic contextualism (hereafter just contextualism). As we saw in the last chapter, contextualism's diagnosis of the skeptical problem turns on positing that "know" is a context-sensitive term, and that one of the major determinants of its use are the conversationally relevant standards. This context-sensitivity, in turn, explains the pull of the ordinary intuition when the standards are typical of everyday use, and the pull of the skeptical intuitions, when higher standards have been put in play by introducing unrealized possibilities of error. We saw also that the primary evidence contextualists advance in favor of this semantic diagnosis of the skeptical problem are salience effects involving "know." Over the course of a number of empirical studies, I showed that salience effects are genuine, and that we do tend to ascribe knowledge differently in different contexts, but that standard versions of contextualism cannot account for our ordinary indifference to implausible error sources.

One of the primary claims I will defend in this chapter is that while the standards for knowledge ascription can vary across contexts, the ultimate explanation of this is not semantic, but cognitive. To present this case, I'll argue that the intuitions behind the skeptical problem have the hallmarks of outputs from domain specific, cognitively impenetrable systems, and not linguistic context sensitivity. But to understand how I'll defend this idea, it's necessary to start by looking at a phenomenon that has been dubbed "semantic blindness," which is at the center of the most significant standing objection to the contextualist's diagnosis.

According to contextualists, if "know" is context-sensitive, then the sentence "I know I have hands" can be true when uttered in a context with low epistemic standards, while the same sentence, in the mouth of the same speaker, can be false in when uttered in a context with high epistemic standards. We have knowledge of the external world, but only if we stay out of the epistemic pressure-cooker of the philosophy seminar room. An important consequence of the view is that these two sentences, even in the mouth of the same speaker, do not contradict each. Their conjunction is not a contradiction because the semantic content of "know" is different in each.

But this resolution of the skeptical problem requires a robust error theory. For, native speakers generally recognize context-sensitive terms for what they are, and can distinguish their contents across contexts. But at the heart of the skeptical paradox is the

intuition of conflict, that the conclusion of the skeptical arguments and ordinary knowledge claims cannot both be true. To explain the phenomena contextualists propose that when it comes to "know," we have a kind of semantic blindness to its context-sensitivity, and this accounts for the intuitive feeling that there is a contradiction.

In what follows I look first at the intuition of conflict and the contextualist explanation of it in terms of semantic blindness. I'll then look at the kinds of objections that have been leveled against contextualism's semantic proposal, especially semantic blindness, and three different types of response that contextualists have offered. In the next section I'll then begin to look at these proposals experimentally. Finally, I'll argue that those studies reveal variation in the grip that the intuition of conflict has on ordinary speakers, and that it is best explained in terms of a conflict between the outputs distinct cognitive systems, not a systematic error like semantic blindness.

2. The Intuition of Conflict and Semantic Blindness

For contextualists to explain the enduring grip of radical skepticism, they must explain why we are adept at making context-sensitive knowledge attributions, but why we are inept at recognizing the logical consequences of those attributions. Recall again the Argument from Ignorance (AI):

Ordinary Intuition	a)	I know I have hands
Error Intuition	b)	I don't know I'm not a BIV
Closure Intuition	c)	If I don't know I'm not a BIV, then I don't know I have hands.
{c.i.}	d)	Therefore, I don't know I have hands
Intuition of conflict	e)	"I know I have hands and I don't know I have hands" is a contradiction.

On the left I've labeled the intuitions that seem to drive AI. According to the standard diagnosis, having each of these intuitions is typical: call this the *philosophical package of intuitions*, or the philosophical package for short. The armchair diagnosis of AI takes it that (a) through (e) are intuitively compelling. The task of offering a response to the skeptical problem revolves around which of (a) through (e) you want to accept, and which you wish to

reject. As Keith DeRose describes AI,

> There are three relevant issues to our puzzle: Is the first premise of AI true? Is the second premise true? Is the conclusion true? And it's easy to endorse the intuitively correct answer to two out of the three questions if you're willing to take the implausible stand on the remaining one.[1]

But the contextualist wants to endorse (a) through (d), and the way they do so is by positing that "know" is context sensitive.[2] This is its semantic thesis, and it involves a kind of context-sensitivity that is familiar. The sentence (1) "Tom is tall" might be true in the context of the elementary class where he teaches, while (2) "Tom is not tall" might be true in the context of the basketball league where he plays after work. Because "tall" is context-sensitive, the semantic contribution it makes to the sentence that embeds it, and so the truth conditions of the whole sentence, can vary with context. Many contextualists, including DeRose, Stewart Cohen, Peter Unger, and David Lewis deploy scalar absolute adjectives like "tall" and "flat" as semantic models for their proposed account of the semantics of "know."[3] Indeed, "tall" and "flat" have been popular, and Cohen outlines the proposal by comparison with similar terms:

> Many, if not most, predicates in natural language are such that the truth-value of sentences containing them depends on contextually determined standards, e.g. 'flat', 'bald', 'rich', 'happy', 'sad'....These are all predicates that can be satisfied to varying degrees and that can also be satisfied simpliciter. So, e.g., we can talk about one surface being flatter than another and we can talk about a surface being flat simpliciter. For predicates of this kind, context will determine the degree to which the predicate must be satisfied in order for the predicate to apply simpliciter. So the context will determine how flat a surface must be in order to be flat.[4]

By positing that "know" is context-sensitive in a similar way, contextualists are able to endorse (a) through (d) in the relevant context, and thereby preserve the intuitions behind the standard diagnosis. The account explains both the appeal of the ordinary intuition and the appeal of the skeptic's argument - and this is an important theoretical virtue of the view.

But this semantic escape route faces two very serious objections. The first is that "know" does not otherwise exhibit the linguistic behavior of uncontroversial context-

[1] K. DeRose, 1995: 28.

[2] Of course they wouldn't endorse them in the same context. In the appropriate context each is true.

[3] P. Unger, 1975; D. Lewis, 1996.

[4] S. Cohen, 1999: 60.

sensitive terms, and this raises the question of whether or not the theory is motivated by linguistic facts, or purely by the solution it offers to the skeptical problem. The second problem is that because contextualists endorse (a) through (d) they take a stand on (e). They argue that the mistake we make is to think that the sentence quoted in (e) is a contradiction, even though it is not, because the "know" in the first conjunct of the sentence makes a different semantic contribution than the "know" in the second conjunct. According to contextualists like Cohen and DeRose, we fail to recognize the context-sensitivity of "know," and this sets us up to be gripped by the skeptical problem. So, it is the intuition of conflict that is really at the heart of the skeptical problem for contextualists. But this means contextualists must posit an error-theory to explain why we make this mistake. I'll consider these problems in turn, and offer my reasons for why we should think the latter is the more damaging.

2.1 Tests for context-sensitivity

So, why think "know," a verb, is a context-sensitive term similar to scalar adjectives? Jason Stanley, for one, argues that it is highly implausible that "know" is a covert, or subtly context-sensitive term, because it doesn't exhibit the kind of linguistic features that most context-sensitive terms do. For instance, according to Stanley one standard test for scalability -also sometimes called "gradeability"- is that scalar terms term take modifiers like "very."[5] So, the following sentences should all sound quite normal.

(3) The table is flat

(4) The table is very flat

(5) Dissertations are hard to write

(6) Dissertations are very hard to write

Moreover, scalar terms allow for comparative constructions, like "taller than," "flatter than," and "richer than." Stanley acknowledges that "know" does accept some modifiers like "really", so both of the following sentences should sound fine as well.

(7) I don't know that the bank will be open

(8) I don't really know that the bank will be open

But as Stanley points out, with scalar terms like "tall," it's possible to conjoin unmodified

[5] J. Stanley, 2004: 124.

assertions with modified negations without contradiction. So the following should sound fine.

(9) Dissertations are hard to write, but not very hard.

While the following pair should sound strange.

*(10) The bank is open, but not really open

(11) He knows the bank is open, but he doesn't really know it.

He points out further that it is usually possible to construct comparisons that exploit the scalar nature of these adjectives. So, the following should sound normal.

(11) Kansas is flatter than Missouri

While a similar construction with "know" sounds quite strange.

*(12) John knows the bank will be open more than Al knows the bank will be open

The awkwardness seems magnified when we consider comparisons to distinct direct objects. Consider the following pair.

(13) John's table is flatter than Al's lawn

*(14) John knows the bank will be open more than Al knows Obama is president.

Sentence (13) sounds natural, and exploits a comparison across a shared scale. In sentence (14), however, it's unclear what the comparison is even of.[6] The evidence is certainly not decisive, but Stanley concludes from these examples and others that,

> If the semantic content of "know" were sensitive to contextually salient standards, and hence linked to a scale of epistemic strength (as "tall" is linked to a scale of height), then we should expect this link to be exploited in a host of different constructions, such as natural comparatives. The fact that we do not see such behavior should make as at the very least suspicious of the claim of such a semantic link.[7]

This consideration, as Stanley pursues it, does limit the *prima facie* appeal of contextualists' semantic thesis. That being said, contextualists are not without direction in exploring a reply.

One direction has been to argue that the context-sensitivity of "know" is derivative of one of its analytic constituents. Cohen applies this strategy by focusing on the justificatory condition of knowledge.

[O]n my view, justification, or having good reasons, is a component of

[6] All the basic types of comparison here are modifications of examples in Stanley (2004: 125 - 129).
[7] J. Stanley, 2004: 130.

> knowledge, and justification and certainly comes in degrees. So context will determine how justified a belief must be in order to be justified simpliciter.[8]

Keith DeRose follows a similar path to the same destination, suggesting that "*S is certain that p*…is a better candidate than the ever-popular *S believes that p* for expressing the attitude requirement for *S knows that p.*"[9] If this is right, and because "certain" *is* a scalar absolute adjective, then the context-sensitivity of knowledge would emerge at an earlier stage. For example, because the attitude ascription in question describes the subjective state of the speaker, not the normative propriety of their stance, it could be a brute fact that in a normal context a person is subjectively certain that they have hands, and that in a skeptical contexts they are not subjectively certain.

In her consideration of Jason Stanley's critique of contextualism, Barbara Partee agrees that the analogy between "know" and "tall" is forced, but she also argues that there are other and better analogues for contextualists to explore. In particular, she thinks the must fruitful possibility for explaining contextual variation in knowledge ascriptions is non-lexical standard-setting, of the kind introduced by David Lewis in his work on scorekeeping and language games,[10] the classic example of which, discussed by both J. L. Austin and David Lewis, concerns describing the shape of France.

(15) If France is hexagonal, then Italy is boot-shaped.

Sentence (15) is a perfectly acceptable, and depends on contextually changing the standards of use, from loose to strict, rather than any genuinely *semantic* feature of "hexagonal." Partee suggests that perhaps the most promising proposal is Lewis' early suggestion that that non-semantic constraints on discourse, that are settled pragmatically, determine the constraints on use. Partee's observation, like Lewis' original proposal, is intriguing but vague. The appeal of the contextualist's semantic account is that it proposes a clear mechanism, and suggests a clear reason why "know" admits standards changing. Both kinds of explanation are absent or undefined in Partee's proposal.

---

[8] S. Cohen, 1999: 60.
[9] DeRose, 2009: 186.
[10] B. Partee, 2004: 157.

Ultimately speaking, though, there are two questions before us, and the search for analogies is only likely to settle one. One question is whether or not "know" is a context-sensitive term. Here the search for analogies along with data from use is as powerful a tool as can be applied toward a solution. A second question is whether or not the purported context-sensitivity of "know" gives rise to the appeal of the skeptical problem. And here the search for analogies is far less likely to be decisive. For, even if Partee and Lewis are right that non-semantic standards of use can change when using "know," as with "hexagonal," it is a further question whether or not the skeptical problem turns on, or in any way involves something like that phenomena. For example, although it's possible to use "hexagonal" loosely to describe France, it by no means follows that any debate over whether something *is* hexagonal depends on such loose use. Some debates might, but some may not. Consider the following dialogue: *Steven: You know, France is really hexagonal. Derek: What!? France is like completely irregular, a hexagon has flat sides! Steven: of course, of course. I just meant loosely like a hexagon, not that it is actually a hexagon.* The dialogue makes sense, even if Derek's misunderstanding is a little forced. But if Steven and Derek disagreed about whether or not the United States military headquarters in Washington D.C. was a pentagon, or a hexagon, it would be a real disagreement, and not one that depended on loose use.

But notice, it seems like the debate over the skeptical problem is more like the latter disagreement and not the former – for the latter is a factual claim, and it seems like one person must be right, and the other wrong. The former disagreement turns on a misunderstanding, and it's easy to interpret both Steven and Derek as having said something true according to the standards they use. Indeed, once Steven explains himself, it should be obvious to everyone involved that he was speaking loosely. A question the contextualist must answer is why the skeptical problem seems to involve something like the factual disagreement rather than the problem of misunderstanding. It's this question, and the most serious objection to the contextualist diagnosis of the skeptical problem that we turn to next.

2.2 Semantic Blindness

The problem, then, for contextualists is to explain why we don't recognize that

"know" is context-sensitive when we generally do recognize other context-sensitive terms as such. For example, "I," uttered by different speakers, Ida and Hazel, makes a different semantic contribution to their respective utterance. Consider the following pair of sentences.

Ida says:

(16) "I'm happy"

Hazel says:

(17) "I'm unhappy"

We all recognize that those sentences mean different things. Their meanings are transparent. We certainly do not think that Hazel is contradicting Ida because we recognize the contextually specified content of the first person pronoun "I" without any conscious reflection. But if "know" is context-sensitive, then our unique blindness to its context-sensitivity does need an explanation. The problem comes to the fore in understanding the contextualist resolution to the skeptical problem. For, as we saw the standard contextualist line is that the intuitive conflict we feel between ordinary knowledge ascriptions and the conclusions of skeptical arguments is a mistake. Here's DeRose,

> What we fail to realize, according to the contextualist solution, is that the skeptic's present denials that we know various things are perfectly compatible with our ordinary claims to know those very propositions. Once we realize this, we can see how both the skeptic's denials of knowledge and our ordinary attributions of knowledge can be correct.[11]

But, we don't tend to make mistakes of the same sort with other context-sensitive terms, so contextualists must posit an error-theory to explain this failure. To see why, we can consider an example from before. I might utter (1) "Tom is tall" in the context of the elementary class where he teaches, and latter the same day, I might say (2) "Tom is not tall" in the context of the basketball league where he plays after work. Having done so, no one who heard both utterances would think that I'm inconsistent, or that I had contradicted myself. We're perfectly sensitive to how context specifies the reference class for "tall," if only tacitly. But contextualists must suppose that we make something like this very mistake to explain why the ordinary assertion "I know I have hands" and the skeptical denial "I don't know I have hands" seem to straightforwardly contradict each other.

[11] K. DeRose, 1995: 5.

The status of the intuition of conflict is the most pressing for the contextualist diagnosis of the skeptical problem. Contextualists aim to motivate their position by appeal to salience effects in low-scrutiny contexts, and not by appeal to skeptical arguments. And as we will see in Chapter 4, this extension is deeply flawed. But even leaving those results aside, whether "know" is a context-sensitive term is one question. Whether we are receptive to skeptical arguments *because* "know" is context-sensitive is a very different one. The former can be true without the latter. So, even if contextualists can offer a reply to the kinds of linguistic disanologies we saw above, the plausibility of their diagnosis of the skeptical problem turns on being able to motivate an explanation of (e) in the standard diagnosis.

Steven Schiffer, among others, has argued that contextualism's error theory is deeply implausible because they must posit a kind of "semantic blindness" to explain why we feel the intuition of conflict.[12] As he frames the problem, the contextualist is committed to "the claim that people uttering certain knowledge sentences in certain contexts systematically confound the propositions their utterances express with the propositions they would express by uttering those sentences in certain other contexts."[13] Schiffer's complaint is that ordinary speakers just don't make this kind of mistake. He argues further that, "that error theory has no plausibility: speakers would know what they were saying if knowledge sentences were indexical in the way the Contextualist requires."[14] Jason Stanley argues in a similar fashion that to preserve their diagnosis of the skeptical problem, contextualists must posit "a hitherto unknown form of linguistic ignorance."[15] The proposal can begin to sound like special pleading. The problem is a serious one for contextualists as there is a deep tension between the ordinary language basis of contextualism's semantic theory, which requires linguistic competence in knowledge ascriptions, and the error-theory, which requires linguistic incompetence in interpretation.

2.3 The Generality Response to Semantic Blindness

Contextualists have offered three kinds of reply to the problem. According to the first semantic blindness is quite general; according to the second it is general among scalar

[12] Other critics of the proposal include: Bach, 2005; Conee, 2005; Feldman, 1999; Hawthorne, 2004; MacFarlane, 2005; Richard, 2004; Rysiew, 2001.
[13] S. Schiffer, 1996: 325.
[14] S. Schiffer, 1996: 328.
[15] J. Stanley, 2005: 116.

adjectives; according to the third semantic blindness is a problem for both contextualists and invariantists. I'll take these proposals up in turn, an offer my reasons for thinking that only the latter two are at all viable.

Martin Montminy argues that semantic blindness is not unique, contra Jason Stanley, and thus it is not an *ad hoc* proposal; call this the generality response.[16] Montminy points to the debate over semantic minimalism as an example. Semantic minimalists like Herman Cappelen and Ernest Lepore argue that genuine semantic context-sensitivity is restricted to a relatively small set of terms, which they call the "basic set", and includes indexicals like "I," demonstratives like "this" and "that," and pronouns like "him," "her," and the like. According to semantic minimalism, most terms make a determinate contribution to the truth-evaluable content of a sentence irrespective of the context in which they are used. But many philosophers disagree, and believe that context plays an essential role in determining content by filling the tacit parameters of many terms. Consider the sentence:

(18) Every box is empty.

According to semantic minimalists, this sentence expresses a single, determinate proposition, *every box is empty* and that it applies to every box in the world, and it is only true if, indeed, every box in the world is empty. It is, in technical terms, semantically complete. Of course we often utter (18) to express the proposition that *every box in the kitchen cupboard is empty* when we are looking at our kitchen cupboard. Cappelan and Lepore argue that context comes into play at the level of speech act, and that the minimal proposition in conjunction with context determines the proposition expressed by the speech act. Many philosophers disagree. Jason Stanley and Zoltan Szabo, for instance, argue that (18) is semantically incomplete because the quantifier "every" has a tacit parameter that must be determined by the context of utterance. According to Stanley and Szabo, (2) does not express a determinate, truth-evaluable proposition until a context is specified.

Leaving aside how the debate has been argued, Montminy's point is simple: either way, somebody is right and somebody is wrong, and whoever is wrong is experiencing semantic blindness regarding the quantifier "every." If the minimalists are right, then it's the contextualists who are blind to some of the features of their language - and that's despite

[16] M. Montminy, 2009.

being quite competent with it. If the contextualists are right, then it's the minimalists who are blind to the workings of their language, again, despite being quite competent with it.

Montminy's generality response seems to point to a different kind of blindness than is implicated in the contextualist's diagnosis. I think Montminy is right that both experts and layman cannot tell by direct introspection whether or not quantifiers have tacit parameters or not, but this is a blindness to the underlying structure of our language, and not one to the content a speaker expresses by using a term. In fact, I think everyone has blindness of this kind - that's why theoretical disputes are argued and decided on the basis of evidence from usage, not introspection. A parallel blindness would be if one person found both the claim that *every box is empty* intuitively compelling (uttered in a context where the living room was salient, and every box was empty), and the claim *not every box is empty*, equally compelling (uttered in a context where the whole world was salient). This kind of conflict turns, as does the contextualist's proposal in the skeptical case, on their having expressed different propositions with the sentences using "every," and then failing to realize the domain over which "every" extended. But that kind of confusion is bizarre, and materially different than what Montminy points to in defense of contextualism's error theory.

2.4 The Close Analogues and Partial Blindness Responses

There are two other more promising lines of response. Stewart Cohen has argued that many of the semantic analogues to "know" that contextualists point to, including "flat" and "tall" *can themselves* lead to a similar kind of semantic blindness. Keith DeRose, by contrast, has intriguingly suggested that perhaps the intuition of conflict is not as widespread as traditionally assumed, and this relieves the contextualist from the burden of having to posit a peculiar error theory in defense of their diagnosis of the skeptical problem.

If Cohen is right, then contexualists can avoid the charge of special pleading. If DeRose is right, then contextualists are not unique in having to posit semantic blindness. Both proposals face two questions. First, and simplest, Cohen and DeRose make different empirical predictions, and we can ask who is correct? Second, and more substantial, we can ask how successful each is as a response to the skeptical problem. As I plan to put each proposal to an empirical test, I'll take these questions up in turn.

3. Testing the Intuition of Conflict & the Close Analogues Response.

Stewart Cohen's proposal is perhaps the most promising defense of contextualism's error theory. He claims that semantic blindness is a more general phenomenon than previously realized, and can be found to crop up with other uncontroversial context-sensitive terms that are close analogues to "know" according to contextualism. Following Peter Unger, Cohen suggests that scalar absolute adjectives[17] can be used to generate skeptical arguments very similar to epistemic ones, and that they will likewise result in a similar intuition of conflict. Unger argued that nothing is truly flat, or that at most there is only one flat thing, because whatever object you take, there is always something that is, or could be flatter. As we saw earlier, many properties measured on a scale with closed upper bounds admit similar treatment: other scalar absolute adjectives include, "empty," "solid," "circular," and "certain." And if, as Cohen argues, "flat" is clearly context-sensitive, semantic blindness is not an *ad hoc* posit, but something to be fully expected with a term like "know" that is perfectible, or allows completeness. In Cohen's own words:

> Unger's case for flatness skepticism is interesting precisely because many who feel the pull of flatness skepticism look back on their previous flatness ascriptions and think they may have been wrong. So the tendency of competent speakers to think that their skeptical inclinations conflict with their everyday knowledge ascriptions does not count against a contextualist interpretation of those ascriptions. *On the contrary, the error theory component of contextualism predicts that competent speakers will think there is a conflict.*[18]

Summing up Cohen's proposal semantic blindness is a typical feature associated with scalar absolute terms: such terms can be used to create skeptical arguments by ratcheting up the standards, and having done so, an intuition of conflict will result.

DeRose offers a different response to the challenge from semantic blindness. He argues that semantic blindness is a partial phenomenon, and that some people have it and some people don't. It's an intriguing proposal because it involves a substantial revision of the standard description of the skeptical problem. According to his account, many people *do* see that knowledge claims are context sensitive when they are presented a properly constructed case, which is fully described and includes facts about the presuppositions,

[17] Cohen himself does not use this terminology, but his selection of examples suggests he has in mind what modern linguistics categorizes as scalar absolute adjectives. You can compare non-absolute scalar adjectives like "tall" or "happy" to absolute ones like "flat" or "certain," where "certain" has a closed upper bound. An agent can be "fully" certain, but not "fully" tall. See, Murphy, 2010: 233.

[18] Cohen, 1999: 82 (emphasis added).

purposes, and the intentions of the conversational participants. He predicts that if you asked an ordinary speaker "if the conclusion of the skeptical argument contradicts what she would say by asserting 'I know I have hands' in an ordinary, non-skeptical context…*many ordinary speakers will say 'yes,' and many will say 'no'.*"[19] It is, according to DeRose, "just a hard question."[20] Indeed, he reports that after this is done, and when he polls his students about whether the knowledge attributing claim and the knowledge denying claim contradict each other "both positive and negative answers tend to be strongly represented."[21]

Before asking how this helps the contextualist, it's worth just registering that by polling his students DeRose seems to be acknowledging this is an empirical question. But in doing so, we should also register that there are serious reasons to be wary of such informal polls. Students are *keenly* aware of the answers their professors *want* them to give. The effect is so profound, that serious experimentalists have to labor very hard to avoid "accommodation" of the same kind among anonymous, uninformed, unbiased survey participants who have nothing to gain from answering one way or another.

But to return to the question at hand, how does this help the contextualist? The idea is that if invariantists are right, and the intuition of conflict is a response to a genuine contradiction, then the proportion of his class that propose the contextualist resolution are experiencing semantic blindness - a fact invariantists will have to explain. If contextualists are right, and the intuition of conflict is a mistake, then the proportion of his class see it as an intuitive contradiction are experiencing semantic blindness. Either way, according to DeRose, "That there is a good deal of 'semantic blindness' going on here is simple a fact that any credible analysis of the situation must face."[22]

These proposals differ in their predictions about the intuition of conflict, and my first goal will be to test these predictions. Cohen predicts that the intuition of conflict is general, and even occurs with other scalar-absolute terms. DeRose argues that it is partial. The difference is not an idle one when it comes to explaining the skeptical problem. If Cohen is right, and the intuition of conflict can be found to occur with other uncontroversially context-sensitive terms, this would lend substantial support to

[19] K. DeRose, 2009: 177.
[20] *Ibid.*
[21] DeRose, 2005: 194.
[22] DeRose, 2009: 178.

contextualism's semantic diagnosis of the skeptical problem. The partial blindness defense is much more philosophically problematic. If contextualists are right about the semantics of "know," and DeRose is right that many people recognize this, it raises the question of why the skeptical problem has exercised such a long-standing grip on philosophers? If many people *don't* experience semantic blindness, wouldn't we have resolved the problem long ago? But this points to a deeper problem for the partial blindness response - positing semantic blindness is not the only explanation of individual differences when it comes to the intuition of conflict. I'll take up this issue at much greater length in section four, but for the time being it is enough to point out that having conflicting intuitions does not entail that the conflict arises because of semantic blindness. The expression *I'm of two minds on the matter* is a commonplace one, and it expresses being pulled in two directions on a given proposition. I think many people were of two minds when it came to the question of whether we should invade Afghanistan. But also many weren't - many people believed firmly in one course of action. But the best explanation of this is not semantic, it is not that "invading Afghanistan is the right thing to do" involves some covert context-sensitivity, to which some are prone, and others not. It is that the conflict of some and the resolution of others depend on deeper commitments. It is possible to be of two minds about a matter, without experiencing semantic blindness. Assume an invariant account of "know," in that case a person could recognize that the two claims contradict each other, but feel intuitively pulled toward accepting both on different grounds. Likewise with the assumption of a contextualist account: a person could experience no semantic conflict, but nevertheless feel intuitively pulled toward rejecting one of the claims on other grounds. That the conflict can lie at a deeper, independent level is an option I'll explore in section four.

3.1 Experiment 1: Empirically Describing the Intuition of Conflict

Cohen's position predicts that most people will feel the intuition of conflict, while DeRose's partial blindness view predicts something like an equal division between those who do and those who don't. Which is right is a question we can settle empirically.

Participants, Materials, and Procedure:

The goal of this study was to have participants affirm a statement, and then motivate its negation through an argument modeled on AI, and then test whether they felt the

intuition of conflict. Participants included 46 college students and 37 college graduates, ages 18 - 65.[23] Probes were administered in a classroom setting and using online survey software. The test involved two independent conditions, each with 41 participants. The control condition, which I call *the fact condition*, presented participants with a clear, non-context sensitive contradiction of the form "P & Not P" to test whether ordinary speakers reason in accord with the law of non-contradiction. The experimental condition, which I call *the skeptical condition*, presented participants with a parallel construction, involving a knowledge affirming clause and a knowledge denying clause of the form "I know that P & I don't know that P" to see if ordinary speakers feel the pull of the intuition of conflict when knowledge attribution is involved.

Each condition involved three stages. The first stage presented participants with a series of true/false statements to determine their beliefs about a number of questions. Question 1 became the basis for the fact condition, and question 2 became the basis for the skeptical condition. The questions, below, were randomized in the survey presented to participants. Questions 1 & 2 were chosen because a large majority of people would be likely to affirm them.

1. **The Declaration of Independence was signed on July 4th, 1776**
2. **I know George Washington was the 1st President**
3. I know I have hands
4. I'll eat supper tonight
5. I know where my car is parked

In the second stage participants read a short story that was intended to cast doubt on the focal statement (either 1 or 2 depending on the condition) from stage one. In this stage participants were asked to read an argument modeled on AI. The arguments concluded that "the Declaration of Independence was not signed on July 4th, 1776" in the fact condition, or "I don't know that George Washington was really the 1st President" in the skeptical condition. The goal of stage 2 was to motivate the denial of a statement the participants affirmed in stage one.

[23] One participant was excluded from the analysis for failing to complete the survey.

The third stage aimed to detect whether participants felt an intuition of conflict between their affirmation in stage one, and the conclusion of the argument denying it in stage two. After participants evaluated their respective arguments, they were then asked "Does the conclusion of this argument show that you were wrong when you said 'I know knew George Washington was the 1st President' ?" / "The Declaration of Independence was signed on July 4th, 1776.' ?" Stage 2 and 3 of both conditions are below.

Skeptical Condition:

> In the Soviet Union, during the height of the Cold War, there was a great deal of state-sponsored propaganda, especially in schools. Much of this propaganda focused on burnishing the image and authority of prominent political figures. As fantastic as it seems, an entire generation of students was raised holding some implausible things. Children were taught that Joseph Stalin invented the polio vaccine and that he had invented penicillin. They were also taught that he had composed the Ode to Joy, Beethoven's 9th symphony. They were further taught that he made great contributions to designing new and more powerful tractors. A whole generation was systematically deceived because a generation of teachers, whether out of fear or some misplaced political fervor, was complicit in the deception. Once the deception became apparent, a whole generation was put in the position of questioning much of what they learned in school, especially about history. These students had false beliefs, not because of stupidity or some failure on their part, but because systematic deception is, perhaps, more plausible than one might ordinarily think.

1. I don't know that I'm not the subject of systematic deception.

2. If I don't know that I'm not the subject of systematic deception, then I don't know that that George Washington was really the 1st President.

3. Therefore, I don't know that George Washington was really the 1st President.

4. Does the conclusion of this argument show that you were wrong when you said you knew "George Washington was the 1st President?" Yes - No

Factual Condition:

> One longstanding misconception historians have tried to clear up is that Independence Day celebrates the signing of the Declaration of Independence. It celebrates the ratification of the document. The signing of the United States Declaration of Independence did not occur on July 4, 1776.

The final language of the document was approved by the Second Continental Congress on that date, it was printed and distributed on July 4 and 5, but the actual signing occurred on August 2, 1776.

1. The Declaration of Independence was signed on August 2, 1776.

2. If the Declaration of Independence was signed on August 2, 1776, then the Declaration of Independence was not signed on July 4th, 1776.

3. Therefore, The Declaration of Independence was not signed on July 4th, 1776.

4. Does the conclusion of this argument show that you were wrong when you said "the Declaration of Independence was signed on July 4th, 1776?" Yes - No

Results:

Of the 82 total participants in the introductory survey, 37 in the fact condition and 40 in the skeptical condition answered "true" to the focal question of interest in their condition. Discussion will be restricted to these participants because it is only among this group that it is possible to see if they believed the conclusion of the argument showed that their initial judgments had been contradicted. Answers are scored binomially. Answering "true": was given a score of 1; Answering "false" was given a score of 0. Scores were then summed and the means are given in Table 1 below, the results are visualized in Figure 1.

Introductory survey	**Score**
I know George Washington was the first President	.90 (37-4)
The Declaration of Independence was signed on July 4th, 1776	.98 (40-1)
Skeptical Scenario	**Score**
Does the conclusion of this argument show that you were wrong when you said you knew George Washington was the 1st President?	.50 (20 - 20)
Ordinary Scenario	**Score**
Does the conclusion of this argument show that you were wrong when you said the Declaration of Independence was signed on July 4th, 1776?	.80 (31 - 6)

Table 1: Study 1 results.

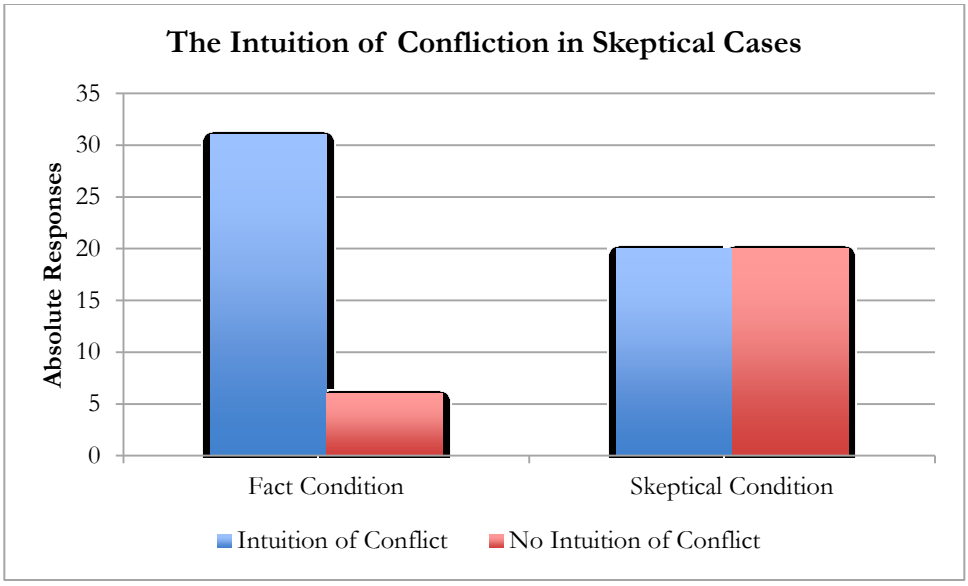


Figure 1: Study 1 results - The intuition of conflict in skeptical scenarios. The fact condition, which served as a control, indicates that non-philosophers do tend to endorse the principle of non-contradiction. The results of the skeptical condition tend to support DeRose's partial intuition of conflict view. A Chi-square test showed there is a significant difference in responses between conditions, X^2=8.36, p<.0034.

Discussion:

In the skeptical scenario, half of the participants did not think the conclusion showed they were wrong previously. In the factual condition, a substantial majority did say that the conclusion of the argument showed that they had been wrong previously. A Chi-Square test showed that there is there is a significant difference in responses between conditions, X^2=8.36, p<.0034, O/R=5.16 (a very large effect). Participants seemed very willing to acknowledge that they had been wrong previously in the factual condition, showing that they do seem to respect non-contradiction, and this is prima facie evidence that they experienced a clear intuition of conflict between the truth of the statement they affirmed in stage one, and the truth of the statement they affirmed in stage 2. The results were mixed in the skeptical condition. About half the subjects said that the conclusion of the argument *does not* show that they were wrong previously. One interpretation of this set of responses is that they did not feel the pull of the intuition of conflict, and this bears out DeRose's prediction that positive and negative answers will be strongly represented.

It is, as I've observed a number of times already, a further inference to conclude this pattern is the product of semantic blindness - an inference I'm as yet unwilling to endorse.

The results also raise a number of questions, both methodological and philosophical. First, using a forced choice question to measure the intuition of conflict doesn't provide much information about the strength of participant's response. Second, although on the face of it the results bear out DeRose's prediction, they of course do not address Cohen's prediction that the intuition of conflict, and thus semantic blindness, arises with other scalar-absolute terms. It's possible that the intuition of conflict is complicated, and that some people have it, some people don't, and that *that* pattern emerges with other context-sensitive terms. If so, and if Cohen's prediction is accurate, it would still provide positive support to the contextualist's semantic diagnosis of the skeptical problem. Pursuing this defense still has its costs, however. As we saw, the standard diagnosis of the skeptical problem rests a great deal of weight on (e) and the intuition of conflict. Bearing out DeRose's prediction is a mixed blessing for the contextualist position. While it does relieve some of the burden of positing semantic blindness, it also raises for them the question of why philosophers seem so predisposed to semantic blindness. For, presumably, the underlying account of *why* philosophers are gripped by the skeptical problem hasn't changed since Descartes time, and so, presumably, the fact that many people *don't* experience the intuition of conflict will not have changed over that time either. If so, the contextualist must explain why their solution didn't emerge earlier. Why is it that philosophers seem so prone to the intuition of conflict and, as contextualists would have it, semantic blindness? Now, as I've already tried to defend, I think that the skeptical problem *is a philosopher's problem,* but not a semantic one; but what contextualists must begin to explain is how it could be that after a century's obsessive focus on language, it's *philosophers* who are disproportionately prone to linguistic incompetence error that leads to skepticism.

But before tackling this question head on, there is another line of response to this study that deserves to be explored. This study did not consider participants' evaluations of the underlying argument prior to assessing the intuition of conflict. It's possible that participants in the skeptical condition were just much less likely to endorse the conclusion because the possibility of systematic deception is a far less plausible source of error than mistaking the ratification of the Declaration of Independence with the date on which it was signed. Perhaps participants in the skeptical condition did not believe the conclusion of the argument showed they were wrong because they did not believe the conclusion of the argument.

3.2 Experiment 2: Testing the Close Analogues View Directly

To explore this possibility, and Cohen's close analogues proposal, I prepared a second study. Recall Cohen's prediction:

> Unger's case for flatness skepticism is interesting precisely because many who feel the pull of flatness skepticism look back on their previous flatness ascriptions and think they may have been wrong…*the error theory component of contextualism predicts that competent speakers will think there is a conflict.*[24]

This defense of the contextualism's error theory rests on two empirical predictions: i) that ordinary speakers will feel the pull of scalar-absolute skepticism (such as "flatness skepticism" – that nothing is truly flat) and ii) once gripped, they will also experience an intuitive conflict between their ordinary perceptual commitments and the conclusions of these skeptical arguments. To be clear, Cohen predicts that once an ordinary speaker is presented with an Unger-style argument they will agree that "nothing is flat." He also predicts that they'll then agree with the sentence "I was wrong to say 'X' is flat." If i) and ii) hold for scalar absolute adjectives, it will not only show that semantic blindness is more general than critics like Schiffer realized, vindicating contextualism's error theory, but it will lend support to the contextualist's basic claim that the skeptical problem turns on the semantics of "know."

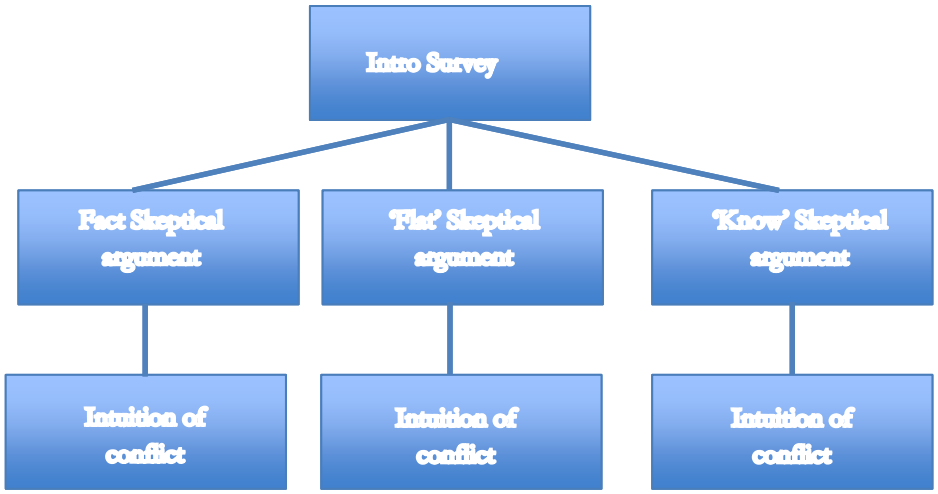
Participants, Materials, and Procedure:

To test these predictions I designed a 3 stage experiment, with 3 conditions, and a between subjects design; see figure 2 which diagrams the experimental design. Participants were 320 (Age, *M*=2.9.5; Gender, Female=45%; Education, 41% Bachelors or above; Race, 17% non-white) individuals recruited on Amazon Mechanical Turk, and compensated at minimum wage for their time. As in study 1, the first stage established participants' pre-theoretic commitments by asking short, true or false questions in an introductory survey. The questions were randomized, and all subjects answered the same questions in the first stage. The second stage introduced three different skeptical scenarios, each of which denied the truth of a commitment introduced in the first stage. Participants received only one of the skeptical scenarios. The first, drawn from the *fact condition* in study 1, argued a widely

[24] Cohen, 1999: 82 (emphasis added).

held historical belief was false. The second, *flatness condition*, argued nothing is flat. The third, *knowledge condition*, was similar to the skeptical condition in study 1, but introduced a global skeptical hypothesis in place of systematic political deception. Each case had the form of AI (see below). Stage two differed from study 1 by asking participants to assess the truth or falsity of the premises and conclusion of the skeptical arguments. Doing so makes it possible to distinguish those who don't have the intuition of conflict from those who don't accept an argument that concludes that one of their basic beliefs is false. The third, and final, stage gauged participants' intuition of conflict. In this study the probe to gauge the intuition of conflict differed from study one. Participants were asked to respond to the following statement on a Likert scale, where 1=Strongly disagree and 7=Strongly agree:

> Given that the conclusion of this argument and the original statement you made say different things, we would like to know whether you think both can be true, or if one must be false. In other words, to what extent do you agree or disagree with the following statement: "My earlier statement that "The Declaration of Independence was signed on July 4th 1776" and the conclusion of this argument which states "The Declaration of Independence was not signed on July 4th 1776" can both be true."



F

Figure 2. Study 2 Experimental design.

Stage 1: Introductory Survey

Participants were instructed to indicate whether they thought each sentence was true (T), or false (F). These questions became the targets for skeptical denial in the second stage.

1. The Declaration of Independence was signed on July 4th, 1776 T – F
2. I know I have hands T – F
3. My kitchen table is flat T – F

Stages 2 & 3: Skeptical Scenarios

Participants were randomly divided into three groups, and presented with a short paragraph (approx. 250 words) motivating a skeptical argument. Participant received only a, b, or c (below). They were instructed to indicate if they thought each sentence in the argument was true or false. Participants were then presented with the intuition of conflict probe

(a) Fact Condition

Below you'll read about a common misconception about American history. After you read the paragraph, you'll be asked to evaluate an argument based on the paragraph. Please indicate whether you think each statement in the argument is true or false. Finally, you'll be asked about how the conclusion of the argument relates to a previous statement you made.

One longstanding misconception historians have tried to clear up is that Independence Day celebrates the signing of the Declaration of Independence. It celebrates the ratification of the document. The signing of the United States Declaration of Independence did not occur on July 4, 1776. The final language of the document was approved by the Second Continental Congress on that date, it was printed and distributed on July 4 and 5, but the actual signing occurred on August 2, 1776.

1. The Declaration of Independence was signed on August 2, 1776 T - F
2. If the Declaration of Independence was signed on August 2, 1776, then the Declaration of Independence was not signed on July 4th, 1776. T - F
3. Therefore, Declaration of Independence was not signed on July 4th, 1776.. T - F

- Given that the conclusion of this argument and the original statement you made say different things, we would like to know whether you think both can be true, or if one must be false.

- In other words, to what extent to you agree or disagree with the following statement:
  My earlier statement that "The Declaration of Independence was signed on July 4th 1776" and the conclusion of this argument which states "The Declaration of Independence was not signed on July 4th 1776" can both be true.

Strongly								Strongly
Disagree	1	2	3	4	5	6	7	Agree

(b) Flat Condition

Below you'll read about the surface properties of objects. After you read the paragraph, you'll be asked to evaluate an argument based on the paragraph. Please indicate whether you think each statement in the argument is true or false. Finally, you'll be asked about how the conclusion of the argument relates to a previous statement you made.

If you think about it, nothing or almost nothing is actually flat. In the case of almost any physical object, you can find another that is flatter. Tables and walls seem flat to the touch, but even when you look at them closely you can see they have textures, ridges, and irregularities. On a microscopic scale these irregularities would be like mountains and valleys. Tables and walls seem flat, but glass is much flatter. Glass seems flat, but actually the surfaces of computer chips are much flatter. Not only that, with every year chips get smaller and flatter. In fact, whatever object you take, there is always something flatter. If you really think about it, nothing is genuinely flat.

1.	No objects are flat	T - F
2.	If no objects are flat, then my kitchen table is not flat	T - F
3.	Therefore, my kitchen table is not flat	T - F

- Given that the conclusion of this argument and the original statement you made say different things, we would like to know whether you think both can be true, or if one must be false.
- In other words, to what extent to you agree or disagree with the following statement:
  My earlier statement that "my kitchen table is flat" and the conclusion of this argument, which states, "my kitchen table is not flat" can both be true.

Strongly								Strongly
Disagree	1	2	3	4	5	6	7	Agree

(c) Know Condition

Below you'll read about a philosophical argument. After you read the paragraph, you'll be asked to evaluate an argument based on the paragraph. Please indicate whether you think each statement in the argument is true or false. Finally, you'll be

asked about how the conclusion of the argument relates to a previous statement you made.

A thought experiment of long-standing interest in philosophy is the idea of a brain in a vat. The idea, common to many science fiction stories, centers on the thought that a mad scientist, machine or other entity might remove a person's brain from the body, suspend it in a vat of life-sustaining fluid, and connect its neurons by wires to a supercomputer which would provide it with electrical impulses identical to those the brain normally receives. The computer would then simulate reality (including responses to the brains own output) and the person with the disembodied brain would continue to have perfectly normal conscious experiences without them being related to objects or events in the real world. Philosophers are interested in these kinds of scenarios because they raise questions about whether we should trust our everyday experiences if we couldn't tell them apart from those in this case.

1.	I don't know I'm not a brain in a vat	T - F
2.	If I don't know I'm not a brain in a vat, then I don't know I have hands.	T - F
3.	Therefore, I don't know I have hands	T - F

- Given that the conclusion of this argument and the original statement you made say different things, we would like to know whether you think both can be true, or if one must be false.
- In other words, to what extent to you agree or disagree with the following statement:
  My earlier statement that "I know I have hands" and the conclusion of this argument which states "I don't know I have hands" can both be true.

Strongly Disagree	1	2	3	4	5	6	7	Strongly Agree

Results:

1st Stage:

Of the 320 subjects who passed the comprehension question, between 93% and 99% indicated that the introductory statements were true. Analyses in the respective conditions are restricted to those who checked "true" to the condition's focal question. This is because only those participants would have their commitments challenged by the skeptical argument in their respective condition.

2nd Stage:

In the fact condition 87% (95/109) judged that all three sentence in the argument was true. In the knowledge condition 53% (48/91 participants) judged that all three sentence in the

argument was true. In the flat condition 81% (90/111) judged that that all three sentence in the argument was true. A Chi-square test of condition by argument evaluation shows a significant variation across conditions, $X^2(2)=34.7$, $p<.0001$, $V=.334$ (a "large" effect).

	Fact	Flat	Know
Premise 1	90%	85%	70%
Premise 2	96%	93%	75%
Conclusion	93%	87%	58%

Table 3: Individual Response Rates to stage 2 skeptical arguments. Percentages indicate the proportion indicating that the statement is "true."

Cohen's first prediction, following Unger, was that scalar absolute adjectives like "flat" could be used to induce skepticism about the flatness of ordinary objects. These results confirm the prediction, as responses to the fact condition and the flat condition were approximately the same – and overall quite high. Quite interestingly, Cohen, like others, assumes that the standard diagnosis is correct, and that non-philosophers will have the standard package of skeptical intuitions. As I hypothesized in the previous chapter, looking at a skeptical argument in detail reveals that there is significant variation in how gripping non-philosophers find the individual components of skeptical arguments like AI. Cohen's proposal seemed to assume that this would parallel Cartesian skepticism about knowledge, but as these results show, people tend to be much less likely to affirm global skeptical arguments.

3rd Stage:

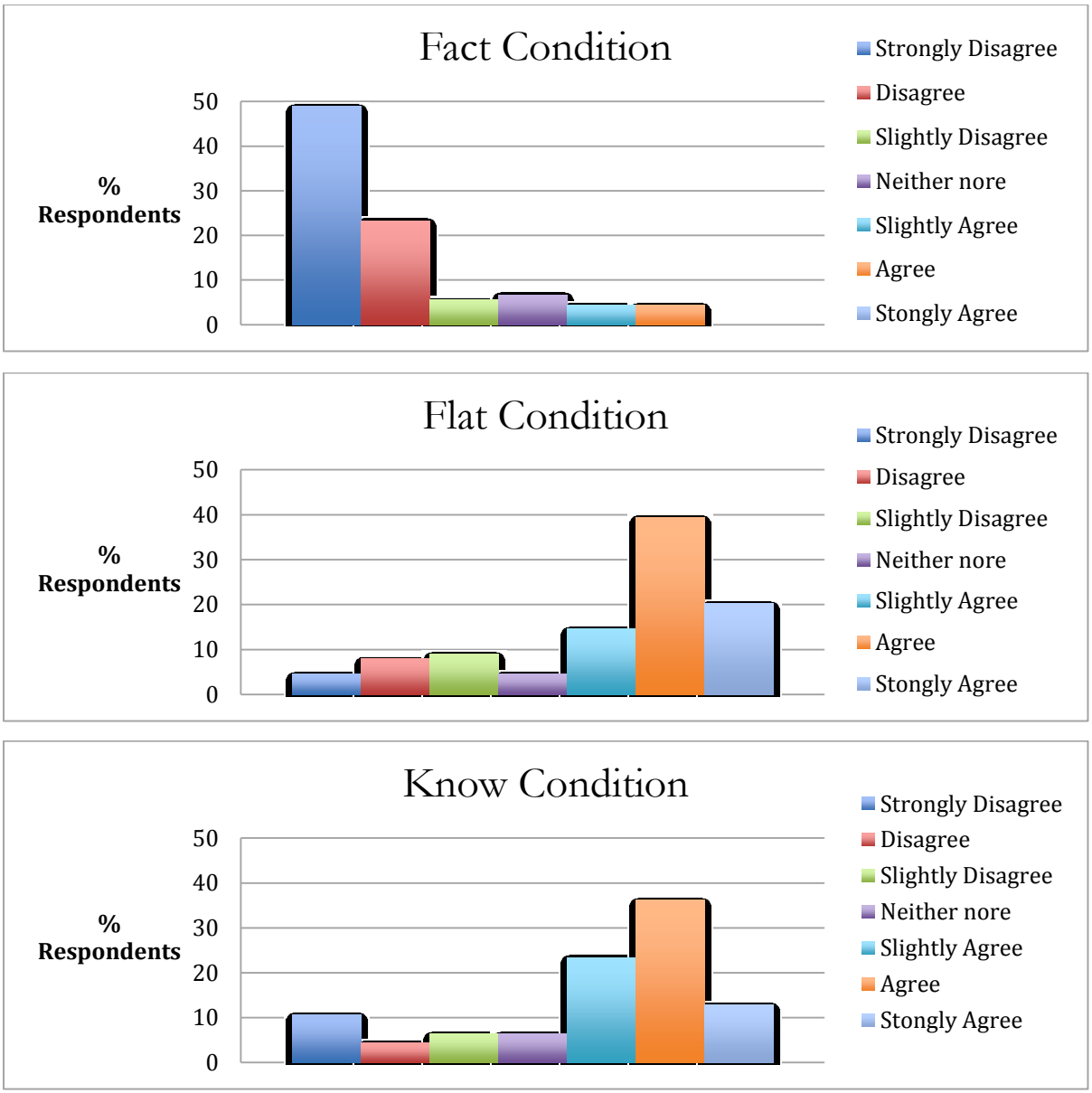
One of the worries prompting this follow-up was that the previous study could not distinguish rejection of a target argument from rejection of the intuition of conflict. Consequently, we'll look at the intuition of conflict among those that accepted their target skeptical argument, and compare it to those who rejected their target argument. The intuition of conflict was judged on a Likert scale, where subjects indicated their agreement/disagreement with the claim that their earlier belief and the conclusion of the skeptical argument could both be true (1= Strongly Disagree & 7 = Strongly Agree). A low number indicates having a strong intuition of conflict, a higher number indicates having no intuition of conflict. The results are summarized in table 4, and visualized in Figures 3-5.

Condition	Accepted 3 Premises	Rejected at least 1 Premise
Fact Condition	M = 2.24 (SD=1.77)	M=3.07 (SD=1.73)
Flat Condition	M = 5.12 (SD=1.76)	M=4.76 (SD=1.51)
Know Condition	M = 4.90 (SD=1.81)	M=3.00 (SD=1.92)

Table 4. Study 2 – Comparing the intuition of conflict between those that accepted the argument and those that rejected at least on premise/conclusion. The results indicate that among those that accepted their focal argument, participants in the fact condition had a strong intuition of conflict as predicted. The also indicate, contrary to Cohen's hypothesis, that they did not have a strong intuition of conflict in the flat condition: participants seemed to recognize that "flat" is context-sensitive. Quite interestingly, participants in the "know" condition who accepted their argument *didn't* have the intuition of conflict! This raises a puzzle for the standard diagnosis *and* contextualists' error theory.

A 2x2 ANOVA using intuition of conflict as a dependent variable and Condition and Soundness judgments as fixed factors, found a main effect for condition, $F(2)=23.4$, $p<.0001$, and near significant effect for soundness, $F(1)=11.01$, $p<.063$. There was also a significant interaction between soundness judgments and condition in determining the strength of the intuition of conflict, $F(2)=9.91$, $p<.0001$.

To untangle these results, I'll begin by looking at the intuition of conflict among those that judged that their respective arguments were sound. A one-way ANOVA confirms what visually inspecting table 4 shows quite clearly: there is a significant difference in intuition between the three conditions, $F(2)=70.04$, $p<.0001$. Follow up T-tests show that the intuition of conflict in the knowledge condition differs significantly from the fact condition, $t(141)=-8.389$, $p<.0001$, Cohen's $d=-1.49$ (a "very large" effect), and likewise between the flat condition and the fact condition, $t(183)=-11.06$, $p<.001$, Cohen's $d=-1.63$ (a "very large" effect), but not between the knowledge and the flat conditions, $t(136)=-.711$, $p=.479$.



Figures 3 - 5: The Intuition of conflict among those who accepted their respective arguments. These results suggest that ordinary speakers detect the context-sensitivity of "know" and "flat".

Looking next at those who rejected at least one statement in their focal argument, a one-way ANOVA confirms there is a significant difference in the intuition of conflict between the three conditions, $F(2)=7.34$, $p<.001$. Follow-up T-tests show that there is no difference in the intuition of conflict between the fact and knows condition, $t(55)=.124$, $p=.902$, but that there are significant differences between the flat and knows condition, $t(62)=3.69$, $p<.0001$, Cohen's $d=1.02$ (a "large" effect) and between the flat and fact conditions, $t(33)=-3.06$, $p<.004$, Cohen's $d=1.05$ (a "large" effect).

What conclusions should we draw from the data? First and foremost, they show that while people do tend to accept the skeptical argument that nothing is really flat (81% acceptance rate), as Cohen predicted, contrary to his prediction they do not have the intuition of conflict with respect to "flat." More interestingly, looking at the knowledge condition, what they show is that those who think the skeptical argument is sound (53% of the sample) do not have the intuition of conflict: they respond just in exactly the same way participants in the flat condition did. But on the other side, those individuals who judge that the skeptical argument is unsound (47% of the sample) do have the intuition of conflict.

Discussion:

Cohen's first prediction was that scalar absolute adjectives like "flat" could induce the same kind of skeptical conclusion that "know" is thought to. These results confirm the prediction. Participants are more likely to accept flatness skepticism than even knowledge skepticism.

However, Cohen's more important second prediction, that scalar adjectives like "flat" generate semantic blindness was not confirmed by these results. Participants did not have the intuition of conflict, and the sensible inference is that this is because the recognized that flatness skeptical arguments introduced a new, more stringent context of evaluation. Agreeing that both statements can be true indicates recognition of the implicit indexing of content to the context of utterance. In the fact condition, a large majority of participants initially agreed the Declaration of Independence was signed on July 4th (98%); of this group, a large majority accepted the subsequent skeptical argument, and they then agreed it was true that the Declaration of Independence was *not* signed on July 4th (87%). This group strongly *disagreed* that both statements could be true (M=2.3); they had a strong intuition of conflict. In the flat condition, while equally large majorities accepted that their kitchen table was flat (93%), and that the subsequent skeptical argument showed that their kitchen table was *not* flat, this group nevertheless *agreed* that both statements could be true because, they did not have the intuition of conflict (M=5.12). This strongly suggests, contra Cohen, that semantic blindness is not a typical byproduct of scalar-absolute adjectives.

The most philosophically interesting results emerge when we look at the knowledge condition. A brief inspection of the distribution of responses in this condition supports DeRose's partial intuition of conflict prediction. Considering the responses dichotomously,

47% of participants indicated having the intuition of conflict by responding below the midline, while 35% indicated they did not have the intuition of conflict by responding above the midline. What DeRose did not predict, and which these results reveal, is that the intuition of conflict seems to be detached from the intuitions driving the rest of the skeptical problem; indeed there is a very strong inverse correlation between judging that the argument is sound and having the intuition of conflict, r=-.457, p<.0001. The 53% who judged that the argument was sound did not experience the intuition of conflict (M=4.9; which is significantly different from the midpoint, $t(47)$=3.42, p<.001). The 47% who judged that the argument was unsound did experience the intuition of conflict (M=3.00; which is significantly different from the midpoint, $t(42)$=-3.43, p<.001). Consider those who accept the skeptical argument, but don't feel the intuition of conflict – for them, on DeRose's account, there is *no* skeptical problem because they recognize the context-sensitivity of "know." But if this is true, DeRose faces the question of why anyone was troubled by skepticism. Why didn't these folks just point it out to the rest of us? Looking at this population, and assuming per partial blindness theory that they have semantic insight, then there should be no skeptical problem. This is a critical question for the contextualist to answer. While most people don't have much in the way of explicit theory, my confident prediction is that if you were to ask the participants in the flat condition *why* both statements could be true, they would say something like "know" just means different things in the statements." And while contextualism has many supporters among philosophers, it is a relatively *new* proposal, and not one that has been backed by appeal to the transparent context-sensitivity of "know." I think the best response available to contextualists would be to accept that there are systematic reasons why philosophers fall into the second group – those who did have the intuition of conflict.

But now focus on those participants who reject the skeptical argument: assuming partial blindness theory again, they think the skeptic's conclusion *would* contradict our ordinary knowledge, but they don't accept the skeptic's conclusion! Again, for this population there is no skeptical problem because they are not attracted to the skeptical argument. These support the partial intuition of conflict interpretation, but they seem to compromise the contextualist's diagnosis of the skeptical problem.

What should we make of these results? Perhaps contextualists can partially square them with the standard diagnosis by distinguishing the intuitiveness of a claim from a

participant's response – we saw an objection to experimental philosophy making use of this distinction in chapter 2. According to this interpretation the participants who did indicate a conflict between the skeptic's conclusion and their ordinary knowledge claim engaged in some naïve philosophizing to avoid the skeptical problem by rejecting part of the skeptic's argument. According to this interpretation, about half of participants have something like the philosophical package of intuitions – they have the ordinary intuition, the intuitions driving the skeptical argument, and the intuition of conflict, but like most philosophers they reject some part of the skeptic's argument as a way to avoid endorsing the skeptic's conclusion.

This would help explain why so few people have the philosophical package of intuitions, despite the assumption that the skeptical problem relies on our ordinary concept of knowledge. But how should the contextualist explain the broad division we see between those that do, and those that do not experience the intuition of conflict? One might typically try to explain systematic difference in the linguistic behavior of different groups by proposing that the two groups have different idiolects, and that those differences explain the differences in behavior. There are many such cases, and idiolects do often trace professional boundaries, offering contextualists a potential explanation of why philosophers behave differently as a group. For example, logical terms like "valid" and "deduction" have technical senses for philosophers that differ dramatically from their colloquial sense, and so philosophers' behavior with those terms tends to differ. However, positing this kind of ambiguity is unpromising because it doubles the kinds of semantic blindness we'd have to attribute. Differences in idiolects are like ambiguities, and tend to be easy to spot. Partial blindness accounts still need to explain whey those that *don't* experience sematic blindness don't just rise up in a chorus and explain what's going on to philosophers. To explain why they don't, contextualists would have to propose something like as follows - *while many speakers do recognize the contextual shifts of "know" across contexts, and thereby don't experience the semantic blindness that gives rise to the intuition of conflict, they <u>are</u> blind to the fact that "know" has a different meaning for philosophers.* This escape route is extraordinarily uncharitable, positing many kinds of error, and it still must attribute to philosophers an exclusive sense of "know" - something we saw that the diagnostic tradition tries to avoid.

So, we're still left with the question, what is the contextualist's best response to these results? It comes in three parts: (1) explain the division in the intuition of conflict by appeal

to partial blindness: some people have the intuition of conflict, others don't; (2) explain the rejection of the skeptical argument by those who do feel the intuition of conflict as naïve philosophizing; (3) posit that there are systematic reasons why philosophers experience semantic blindness. (1) allows contextualists to explain the observed group division with respect to the intuition of conflict. (2) seems plausible, and helps the contextualist explain why so few people responded in the way predicted by the standard diagnosis. Of those who do experience the intuition of conflict, many feel the need for consistency, and reject some part of the skeptic's argument. The contextualist is forced to accept something like (3) to explain the historic interest philosophers have had in radical skepticism. But it is (3) that strikes me as extraordinarily implausible. Academic philosophy over the last century and more can be characterized as having been *obsessively* focused on language. Moreover, many of the most influential Anglophone philosophers of the period owed their success in no small part to an extraordinary sense of the subtleties of English. It strains credulity to suppose that, as a group, philosophers have a systematic *deficit* in their understanding of the English language. While it's certainly true that contextualist solutions to the skeptical problem have been advanced by many philosophers over that time, those proposals involved many different linguistic models, and the best explanation of this is that they were not based upon linguistic insight into the semantics of "know," but rational reconstructions of the conventional wisdom on skepticism. For instance, Norman Malcolm proposed that "know" is ambiguous between a weak sense and a stringent sense; more recently, Rene van Wouden has proposed that "know" is polysemous and has a dozen or more distinct senses; Cohen and DeRose, of course, defend scalar-absolute models of "know."[25] This diversity, along with the many smaller variations in the literature, makes it unlikely that the proposals are based on semantic insight.

Partial blindness is in many ways a much more costly error theory than contextualists were saddled with previously. Before they posited a systematic error, one that implicated all speakers. At least according to this original error theory, there was at least no question of why it gripped philosophers. Now, contextualists must still posit that roughly half the population of fluent, native speakers is prone to a systematic linguistic error, and they must offer an explanation of why most, if not all, philosophers fall into the incompetent half. The

[25] See N. Malcolm (1952) and R. van Woudenberg (2005).

kind of linguistic error they posit is still, to parrot Stanley, hitherto unknown, and one which disproportionately affects otherwise extraordinarily fluent and attentive people.

We need a better explanation of the intuition of conflict than semantic contextualism has the resources to provide. And perhaps one is available. While systematic linguistic errors like contextualism's partial semantic blindness are unknown, it is not uncommon to find systematic individual differences in cognition. In the next section I will argue that a much better explanation of the intuition of conflict is available by looking at system conflicts in cognition.

4. Comparing Linguistic and Cognitive Accounts of the Intuition of Conflict.

The goal of the foregoing studies has been two-fold. The first was to distinguish and test the two leading contextualist responses to the problem of semantic blindness. I argued that the results count heavily against the most promising of the two, Cohen's close-analogues account, and favor DeRose's partial blindness theory. I argued further that the partial blindness theory leaves contextualists with an even more implausible error theory than before. The second was to get an overall view of how non-philosophers respond to skeptical arguments as a whole. Looking at their responses, we see that many people do not have the philosopher's response to the skeptical problem.

Now, these results should be interpreted cautiously, if only because the skeptical scenarios I used are familiar to many through movies like the Matrix and books like *Breakfast of Champions*. It is reasonable to worry that some participants' responses reflect naïve philosophizing, rather than their raw intuitions about (a) - (e) of the skeptical problem. One way to respond to this worry is with further experimentation that looks at the component intuitions in isolation, as we did in examining salience effects in chapter 3. Doing so will allow us to get a better sense of how powerful the component intuitions of the skeptical problem are, without introducing the extra baggage of previous associations. So, in one respect an important result of these studies is just that more, and more focused empirical work on the skeptical problem is warranted, especially work focused on the intuition of conflict detached from familiar Cartesian scenarios. But how should we interpret the intuition of conflict results from these studies? In this section I'll argue that they exhibit the character of cognitive conflicts much more than linguistic errors.

The core argument from DeRose's partial blindness defense aims to relieve the explanatory burden of positing semantic blindness. He predicts that many people do feel the intuition of conflict, and that many people don't. If contextualists are right about the semantics of "know", those who have the intuition of conflict suffer from semantic blindness. If invariantists are right about the semantics of "know," those who don't have the intuition of conflict suffer from semantic blindness. Both theories, according to his account, have to explain it, so contextualists are not guilty of special pleading. He concludes that on either a contextualist or an invariantists' account of knowledge, there "is a good deal of 'semantic blindness' going on here."[26]

But this simply doesn't follow from the presence of conflicting intuitions. Finding two conflicting intuitions equally compelling, as in the intuition of conflict, is a phenomenon that can be found throughout the literature on heuristics and biases, and does not involve semantic processing, nor does it necessarily involve predicating systematic error. In what follows, the major claim I want to defend is that the intuition of conflict is better explained at the psychological level, not the semantic level, because the characteristics it exhibits are typical of cognitive systems conflicts. In particular, I will argue that the observed *individual differences* in the intuition of conflict, its *resistance to insight*, and its *resistance to alteration* are all quite typical byproducts of cognitive systems conflicts, not semantic error. A virtue of the cognitive psychological diagnosis I offer here is that it can accommodate the observed differences in knowledge ascription practices that motivate contextualism, without resorting to a highly implausible error theory.

4.1 Dual Process Theory & Intransigence

One of the key insights to emerge from cognitive psychology over the last thirty years is that cognitive processing - how we think about a question or problem - does not depend on a uniform strategy. Instead, problem solving often involves employing a number of different strategies in a way that is both highly sensitive to context, and which can differ widely across individuals. One key factor is motivation: people tend to selectively allocate their cognitive resources depending on context, and motivation can lead to different

[26] K. DeRose, 2009: 178.

responses to the same information.[27] Experimentally manipulating factors that affect motivation can lead individuals to analyze the same information in very different ways.[28] Factors that affect motivation include the relevance of information to an individual,[29] individual differences in the need to justify themselves to others,[30] and the consistency of the new information with past information.[31] Another key factor is presentation. A key result of the heuristics and biases research tradition is that how information is presented can have a significant influence on the strategy an individual deploys to process it - an example I'll discus in more detail shortly is that small differences in how probabilities are presented, as frequencies or as decimals, can have a large impact on how we reason about them.[32]

The resulting picture of our cognitive architecture goes by the name of dual systems theory, and - as the name suggests - the organizing idea of the theory is that we have two distinct ways of processing information.[33] There is now a broad consensus in cognitive psychology that there are two types of systems for processing information that often run in parallel, and are capable of yielding different conclusions.[34] System two is a slow, consciously directed, sequential, high-effort system. System two is often what we associate with our conscious stream of thought. System one is low-effort, fast, automatic, domain-specific and not open to conscious inspection. System one is thought by many to be highly-parallel, involving a number of "modules" or "heuristics" that produce responses to highly-specific inputs.[35] Like any theoretical framework many aspects of dual systems theory are highly controversial, including the number of modules,[36] whether and to what extent those modules/heuristics are innate,[37] and even whether the division into two broad categories is adequate.[38] Nevertheless, dual systems theory is the operating framework for most cognitive psychology, and two broad experimental results remain well-attested. First, there are

[27] P. Ditto & D. Lopez, 1992.
[28] S. Fiske & S. Neuberg, 1990.
[29] Borgida, & Howard-Pitney, 1983.
[30] P. Tetlock, 1985.
[31] Hilton, Kein, & von Hippel, 1991.
[32] A. Tversyky & D. Kahneman, 1974; G. Gigerenzer, & R. Selten, 1999 & 2001.
[33] J. Evans, 2009.
[34] S. Chaiken & Y. Trope, 1999; J. Haidt, 2001.
[35] S. Pinker, 1997.
[36] J. Fodor, 2001.
[37] P. Kitcher, 1985; J. Tooby, & L. Cosmides, 2005; D. Buller, 2006.
[38] J. Evans, 2009; K. Stanovich, 2009.

individual differences in people's tendency to engage in high or low effort processing.[39] The contextual cues that prompt high or low effort processing affect individuals differently, in part because there are natural differences in individuals' tendency to engage in high-effort processing. For example, given the seemingly simple math question, *if a bat and a ball together cost $1.10, and the bat costs a dollar more than the ball, how much does the ball cost?,* many people intuitively jump to the answer $0.10, but of course the answer is $0.05, as the bat must be $1.05 for the two to total $1.10. Almost everyone feels the *pull* of the first answer ($0.10), and almost everyone *can* answer the question correctly ($0.05) if they're warned that it's tricky, but very few people tend to *resist* the intuitive first answer and engage in high-effort processing of their own accord.[40]

This question, from what's known as the cognitive reflection test (CRT), is one where people exhibit individual differences in their tendency to respond. In this way the division in responses the CRT elicits resembles the split we observed in the intuition of conflict: there are group-level tendencies to respond in one way or another. But the intuition driving the incorrect answer is generally transparent or pliable: people tend to recognize the right answer on their own, when they're promoted to try again. In that way, while there is a population-level difference in intuitive response, those intuitions converge readily on a single answer. The typical explanation of this is that a low-level, heuristic process is behind the intuition that the answer is $0.10, but that when prompted, we all can apply general intelligence to override this response and arrive at the correct answer.

A second well-attested result from the dual systems research program is that the tendencies to use a particular strategy in a particular context, and thus the intuitions that attend to them, are highly intransigent - they are very resistant to insight (introspective access) and they are very resistant to conscious alteration.[41] The CRT test question illustrates the difference: it is introspectively accessible, once the right answer is pointed out, most people acknowledge it, but the cognitive style that gives rise to the mistake is very hard to change consciously. Even when people know they make mistakes of this kind, they find it very hard to change their cognitive style.

[39] K. Stanovich, 2010; S. Fredericks, 2005.
[40] S. Fredericks, 2005.
[41] H. Arkes, 1991; R. Larick, 2004.

There are a number of examples from the heuristics and biases research tradition that exhibit this kind of resistance to insight and alteration. One example is base-rate neglect. What is sometimes called the base rate fallacy is a general tendency, first described by Amos Tversky and Daniel Kahneman, to ignore information about a distribution when reasoning about specific cases.[42] Consider the following case:

> Suppose that the breathalyzers used by police have a false positive rate of 2%, i.e. approximately two percent of the people who fail the test have not been drinking, but that the test is 100% accurate at detecting people who have been drinking. Assume also that only one in a thousand people are actually driving under the influence. Finally, suppose the police stop a person at random, and that person fails the breathalyzer. What is the probability that they are genuinely drunk?[43]

Most people, when asked the question, given an answer around 98%, which is clearly taken from the false-positive rate - although they sometimes modify it slightly, subtracting a percent or so, to accommodate the base rate. The actual answer is close to 5%. It's easy to explain intuitively: because actual drunk people are really rare, testing a thousand people will yield one positive result, but because the failure rate is rather high by comparison, testing at random will result in approximately twenty false positives. Only one of twenty people who test positive is actually positive.

The precise probability can be calculated using Bayes's theorem, where *drunk=actually being drunk, sober=actually being sober, and FT=fail the test*) we know that $p(drunk/FT) = p(FT/drunk)*p(\text{drunk})/p(FT)$. The value of $p(FT)$ is given by the formula $p(FT) = p(FT/\text{drunk})*p(\text{drunk}) + p(FT/\text{sober})*p(\text{sober})$, where $p(\text{drunk})=0.001$, $p(\text{sober})=0.999$, $p(FT/\text{drunk})=1.00$, and $p(FT/\text{sober})=0.02$. Working this all through, we see that $p(FT)=.02198$, and that thereby the probability that a person is actually drunk after failing the test, $p(drunk/FT)= 0.0455$.

Insight, here, is highly context dependent. Many people understand and accept the intuitive explanation I gave first, and fail to understand the formal explanation that follows. Some people, of course, follow the formal explanation relatively easily. Gerd Gigerenzer has found that whether or not people recognize and accept the correct answer and the rationale for it depends on how the information is explained - indeed he's found that a frequentist

[42] A. Tversky & D. Kahneman, 1973, 1985.

[43] This example is adapted from Kahneman & Tversky's work (1973).

presentation can be very helpful. This is important, because it shows that how we process the very same information is highly dependent of presentation: it can determine the response we give and the strength with which we feel it. But even for those that follow both explanations, base-rate neglect is extraordinarily intransigent to alteration. As Amos Tversky and Daniel Kahneman found in their initial studies of the phenomenon, even individuals with exemplary talents and a natural tendency to engage in high-effort cognition tend to both make the mistake and continue to make it even after it has been explained. They found that Harvard medical students and Stanford graduate students in statistics are as prone, or nearly as prone, to base-rate neglect as the general population. Only fairly intensive focus on the fallacy, particularly paired with Gigerenzer's frequentist presentation strategy, seems effective at altering responses.[44] All in all, two decades of research into human reasoning has revealed that there is a long and diverse set of cognitive biases, and while these features certainly do not all appear as a package, intuitional conflicts often exhibit i) individual difference, ii) resistance to insight, and iii) resistance to alteration.[45,46]

[44] A. Tversky & D. Kahneman, 1974; G. Gigerenzer, 1999.

[45] There are many surveys of the literature. Kahneman's anthology (1982) *Judgment Under Uncertainty: Heuristics and Biases* was an agenda-setting introduction. Edward Stein's (1999) *Without Good Reason* considers the implications of the program for philosophy and the study of rationality; Daniel Kahneman's more popular (2013) *Thinking Fast and Slow* offers an up-to-date survey of the field.

[46] By way of example, the Wason-selection task is one of the most thoroughly studied tasks in psychology - it turns on failure to understand the structure of the material conditional. Intransigence has been widely studied in the Wason selection task, and this can serve as a model for testing the empirical adequacy of the contextualist error-theory. The Wason selection task is designed to study the typical person's grasp of the material conditional. Subjects are asked to evaluate whether a rule like the following is true: if a card has a vowel on the front (antecedent), then it has an even number on the back (consequent). Subjects are presented with four cards, and told that there are letters on one side of a card, and numbers on the other. They are then presented with four cards with the following values facing up: A; 7; N; 2. They are then asked which ones need to be turned over to test whether the rule is true. The correct answer is A (P) and 7 (¬Q), because only these could falsify the rule. Typically, only about 5% of subjects choose it (Johnson-Laird & Wason, 1970). The most common answer is A (P) and 2 (Q), at around 45% of subjects (*ibid*). The most popular answer after this is A (P) alone at around 30% of subjects (*ibid*). A number of follow up studies have attempted to assess the efficacy of tutoring on subjects performance. Wason (1969) and Wason and Shapiro (1971) conducted studies where subjects where taught the logical structure of the conditional, and where subjects were instructed to think about each card individually and whether turning it over would prove the statement false. They found that these methods did diminish the rate at which subjects made the most common mistakes, but they did not increase the rate at which they choose the correct cards. In a follow up, Cheng, Holyoak, Nisbet and Oliver (1986) found that a semester long course in logic only reduced the rate at which subjects made at least one error from 75% to 72%.

4.2 Heuristics, Biases, and Widespread Error

It might be tempting to conclude from all this that we are, simply, irrational. But that's too fast. Human reasoning does diverge from the canons of normative theory - deductive logic and probability theory - but that's something we always knew. The insight of the heuristics and biases tradition is that human reasoning diverges from these canons in systematic ways that are sensitive to context, sensitive to the presentation of information, and which vary at the population level. The heuristics we've looked at so far are driven by heuristic system one cognition, which generates responses in certain domains that conflict with the outputs of conscious system two cognition. While many of the most well-known examples that have emerged from cognitive psychology have this structure, many have a different structure, as between two different lower-level heuristic processes that are triggered by different contexts.

By way of considering one example, and its implications for our understanding of the skeptical problem and the intuition of conflict, I'd like to return to Philippa Foot's Trolley problems - an example we first discussed in Chapter 1. The dilemma posed by the trolley problems is a familiar one, but I'll rehearse it briefly. In the *switch scenario* we are to imagine that a trolley is running out of control down a track, and will kill five workers unless we pull a switch to divert it onto a spur track. There is, however, one worker on the spur, and they will be killed if we pull the lever. We are then asked (19) "Is it permissible to trade one life for five?". Many people have the strong intuition that the answer is "yes." Empirical studies, both in the industrial West and in the global South show that approximately 80% of people feel this intuition, and feel it quite strongly.[47] In the *bridge scenario* we are to again imagine that a trolley is running out of control down a track towards five workers, but in this case we are standing on a bridge over the track, next to a large man wearing a very large backpack. The only way to save the five workers is to push the man off the bridge and onto the track in order to derail the trolley. The man, of course, will be killed if we push him. We are then asked the same question, "Is it permissible to trade one life for five?". In this case many people have the reverse intuition, and answer "no." The same proportion of people who think it *is* permissible to pull the switch in the first scenario think it *isn't* permissible to push the man in the bridge scenario.

[47] M. Hauser, et al., 2007.

Although it's an empirical question - and not one I've attempted to answer (yet)[48] - my strong suspicion is that if you were to ask a non-philosopher if their conclusion in switch scenario contradicts what she would say by asserting "It is impermissible to trade one life for five" in the context of the bridge scenario, *many will say "yes," and many will say "no."* It is, to my mind, just a hard question. But this doesn't mean that a good deal, or any semantic blindness going on here. Many people *do* have a strong intuition that our responses to these questions should be consistent - they have the intuition of conflict. Many people *don't* have a strong intuition of conflict, and are happy to allow that the claim is true in the switch case, but not the bridge case. How should we resolve this puzzle?

One possible way to resolve the intuitive conflict for those that feel it is to follow the contextualist strategy. We could say that "permissible" is a covertly context-sensitive term: it's meaning is partly determined by the context of utterance, and it expresses different contents in the different contexts of the trolley problem. But like "know", "permissible", lacks the classic features of context-sensitive terms, and so this proposal invites all the same the problematic linguistic disanalogies that confront the contextualist's position on "know."[49]

But in any case, there is already a better account of the conflict on hand that doesn't court a problematic semantic error theory. The trolley problem is the fruit fly of moral psychology, and converging lines of evidence from many different fields shows that thinking about the switch scenario is psychologically and neurophysiologically distinct from thinking about the bridge scenario. In a landmark imagining study, which has been widely replicated, Joshua Greene and his collaborators showed that when we consider the bridge condition,

---

[48] But conducting such a study wouldn't tell us much about the skeptical case except that there is a another, similar conflict of intuitions. I do, of course wish to argue by analogy here, but experimentally demonstrating that there is *one* more analogy among the many to be found in the heuristics and biases literature doesn't strengthen the case in a material way.

[49] Perhaps there is an alternative. Steven Gross has suggested to me that the question in the trolley case involves elliptical indexing to the content. So, it is possible that the question is interpreted as expressing the proposition "Is it permissible to trade one life for five (in this situation)." Because "this" is clearly a context-sensitive term, and if people tend to believe that what is permissible depends on particulars about the case, then one might not observe the intuition of conflict. If people think that what is permissible does not turn on the particulars of the case, then one might observe the intuition of conflict. While the intuition of conflict has not been directly investigated here, some recent results shed light on the possibility. In a forthcoming paper Eric Schwitzgebel and Fiery Cushman (forthcoming a) found substantial order effects in both philosophers and non-philosophers' responses to the trolley problem. Which scenario an individual considers first affects their responses to subsequent scenarios: individuals have a tendency to try to reply consistently. This tendency is even stronger among professional philosophers. This suggests that individuals do feel a conflict in replying differently to the scenarios, and try to resolve it by answering consistently. That philosophers are even more prone to the need for consistency is a very suggestive result. It might indicate a reason why they are more prone to feel the need to resolve the intuition of conflict in the skeptical case.

and we think about actually pushing a living, breathing person to their death to save the workers on the track, there is a significant amount of activity in our anterior cingulate gyrus, an area associated with emotional engagement. But when we think about the switch scenario, and we think about reaching out to pull a cool metal switch, there is very little activity in the anterior cingulate, or any other areas classically associated with emotional engagement. Greene and his collaborators conclude that:

> Why is it acceptable to sacrifice one person to save five others in the [switch] dilemma but not the [bridge] dilemma?...How do people manage to conclude that it is acceptable to sacrifice one for the sake of five in one case but not the other? We maintain that emotional response is likely to be the crucial difference between these two cases.[50]

Trolleyologists confront the same kind of problem as the diagnostician of the skeptical problem. They aim at offering an account of our intuitions, that not only resolves the tension between them, but explains *why* those intuitions are so forceful. Greene et al.'s proposal is that the conflict in intuitions is a byproduct of those intuitions being produced by distinct psychological systems. On the proposal that Greene now favors, our conscious reflective system (system 2 in the language of dual process theory) is triggered when questions of general welfare are in the offing, as in the switch case. But when physical harm and physical contact are salient, as in the bridge scenario, a lower-level, domain-specific system is triggered, which produces a negative affective response intended by evolution to help us avoid dangerous conflicts.[51] As with other examples in the heuristics and biases tradition, the intuitions that give rise to the trolley problem resist insight and resist alteration. The origin of the intuition is difficult to recognize through introspection - indeed, as Greene points out, the intuition driving responses in the bridge scenario is traditionally associated with *deontology* and rationalism, not emotion. And the intuitions are very hard to alter - although it is certainly possible to upon reflection endorse a particular moral theory, and then endorse the appropriate consistent pair of intuitions, the pull of the component intuitions seems nevertheless to remain quite powerful.

So, the diagnosis that Greene offers of the Trolley problem is really quite simple: we have different systems for solving moral dilemmas, those systems activate in response to different contextual cues, and because those systems process information differently, they

[50] J. Greene et al., 2001: 2107.
[51] J. Greene, 2013: 211-253.

can lead us to have conflicting intuitions toward the same proposition. Of course a natural question to ask is why suppose we have psychological structures that lead to such inconsistencies. Wouldn't this simply lead to incoherence, in thought and in behavior, and doesn't that count against the proposal? The stance that dual systems theorists often adopt in answering *why* questions such as these is to think in evolutionary terms, and it is an illuminating perspective in the case of morality. Greene proposes that we tend to apply a utilitarian standard of promoting the greatest welfare for the greatest number when considering moral dilemmas in the abstract, and when they do not pose any particular harm to us individually or to our genetic relations. This seems to be the standard in play when considering the switch dilemma. But when the context of a moral dilemma introduces a cue that triggers our concern for harm to ourselves, to a genetic relation, or to a friend, we then tend to adopt a standard aimed at avoiding harm. Because the bridge scenario involves close physical contact, Greene proposes that strong emotional responses are triggered to lead us away from acts that could endanger ourselves. Certainly trying to push someone of a bridge could lead to a dangerous struggle. This set of standards makes sense from an evolutionary perspective: we should aim to promote the general welfare of the group, so long as it doesn't pose an adaptive cost to us. Although this could lead to inconsistency and confusion, in thought and behavior when the contextual cues are muted or mixed, typically it will render a straightforward result. It is not surprising, then, that philosophical thought experiments often zero in on just those cases or pairs of cases where inconsistencies crop up.

The overall picture that emerges is one that offers a much more attractive way to account for the intuition of conflict than linguistic contextualism's semantic error theory. The first virtue of this approach is that the observed individual differences and intransigence of the intuition of conflict don't need to be explained away as they have analogues among other known cognitive system conflicts. By contrast, a semantic error theory must explain these away. Second, a psychologically grounded explanation of the intuition of conflict does not involve attributing systematic error to those confronting skeptical arguments. Again, by contrast, a semantic error theory must attribute widespread error. Looking first at the example of the trolley problem, while the systems do lead to conflicting responses to the same proposition, they do not lead to behavioral incoherence; indeed they produce a highly coherent and sensible set of behaviors when viewed from an evolutionary perspective. When considered in the terms of natural selection there is no error in either of these intuitive

responses: on the whole we aim to promote the greatest good for the greatest number, so long as it doesn't involve any sacrifice on the part of ourselves or our kin. Moral theory may aim at consistent impartiality, but mother nature does not.

However, it might seem that some of the earlier examples of statistical reasoning we considered *do* involve attributing systematic errors. For example, isn't it just a feature of the heuristics and biases tradition that we have some quick and dirty heuristics that lead us to the *wrong* answer, ones that we can only recognize as mistaken by using controlled, conscious cognition? This is of course right - people do make systematic mistakes from the perspective of *normative theory* when reasoning about some statistical and logical problems. Those normative mistakes, however, often make sense from the perspective of *evolutionary theory*. Human cognition is the product of an adaptive cost-benefit analysis. And while it would be very nice to be right all the time, the adaptive benefit of being right all the time must be balanced against the investment of time, effort, and cognitive machinery needed to achieve that result. Each of these factors imposes an adaptive cost on the individual: time and effort impose opportunity costs, while cognitive machinery, and the actual grey matter to support it, impose actual metabolic costs. A key prediction of the heuristics and biases tradition is that heuristics can often produce results that approach the accuracy of high-effort cognition, but at a much lower cost, when applied in narrow domains.[52] The result is a framework for understanding seeming conflicts in cognition and behavior that doesn't default to attributing widespread error or irrationality. So far I haven't proposed a full explanation of the intuition appeal of the components of the intuition of conflict - I will offer more on that topic in Chapter 5 - but what's worth noticing here is that the compositional approach to cognition adopted by the heuristics and biases tradition offers a way to explain intuitional conflicts in terms of adaptive pressures that gave rise to those systems.

Where does this leave us when it comes to semantic blindness and the intuition of conflict? So far we've seen that the only empirically tenable account, DeRose's partial blindness theory, must still posit a kind of linguistic error that is otherwise unknown: systematic partial blindness to context-sensitivity. Moreover, the partial blindness account attributes an empirically unexplained kind of error. While it's necessary for contextualists to

[52] C. Cherniak, 1991; G. Gigerenzer & R. Selten, 2002; G. Gigerenzer, 2002.

posit partial blindness to explain the grip of the skeptical problem, the error itself is theoretically inert: whatever kind of mistake it is—infelicity, irrationality, or performance error—contextualism provides no account of *why* it might occur. The error explains the skeptical problem, but the error itself is unexplained. By contrast, once we look at the psychological level its easier to achieve explanatory coherence. The same features that are without precedent in linguistics—intransigence and individual difference—are expected byproducts of cognitive systems conflicts. Furthermore, treating the conflict at the psychological level promises a way to explain the intuition of conflict, rather than treating it like a brute fact. But the promise of an explanation and an explanation are two different things, and so far I haven't proposed a full explanation of the error intuition or the ordinary intuition. That will be the work of the next chapter.

5. Conclusion

In this chapter our examination of the objections to contextualism has gone through many stages. We began with the observation that there seems to be a strong intuition supporting the idea that we have knowledge of the external world, and a strong intuition supporting the skeptical argument denying this. These two intuitions seem to be in conflict, but the contextualist diagnosis of the skeptical problem predicts that they are compatible. So, contextualists posit semantic blindness to explain the intuition of a conflict. We then looked at the two leading attempts to defend semantic blindness, Stewart Cohen's argument that there are analogues to semantic blindness among other scalar adjectives, and Keith DeRose's burden shifting argument for partial blindness. We then tested the empirical predictions of these proposals, finding that i) acknowledged scalar adjectives do give rise to contextual tightening of standards, but not to semantic blindness. We also found ii) that the intuition of conflict is not general, and that while some people have a strong intuition of conflict, others don't. Upon first inspection, these results seem to fit best with DeRose's partial blindness theory. However, we also found iii) that those who do have the intuition of conflict seem to reject some part of the skeptical argument, while those that don't tend to accept the skeptic's argument. These results complicate what I called the standard diagnosis of the skeptical problem. While partial semantic blindness is consistent with this results, I then argued that it is nevertheless more problematic as an error theory than complete semantic blindness. Partial semantic blindness, like complete semantic blindness, is without

precedent among linguistic errors, and appears ad hoc. But this problem is compounded by the fact that partial blindness theorists must further explain why philosophers are systematically among the group afflicted by it.

DeRose's goal, of course, was to balance the burden of explanation between invariantists and contextualists, and it was based on a genuine insight into the intuition of conflict - an insight that more careful empirical study has supported. I then argued that this burden-shifting maneuver is only warranted if there is no better explanation of the observed pattern behind the intuition of conflict than semantic blindness. I concluded this chapter by arguing that there is. Cognitive system conflicts exhibit many of the features we see in the intuition of conflict, and this suggests that a more promising route to understanding the skeptical problem is exploring the psychological foundations of knowledge ascription. I have not offered a psychological account of the skeptical problem as of yet, and my main goal in this chapter was to establish that an empirical view of the problem shows that it is unlikely to be a semantic phenomena. In the next chapter I will present an empirical description of the error intuition, and in the final chapter I will propose and defend just such a psychological account of the skeptical problem.

Chapter 4

Salience, Plausibility, & Skepticism

1. Introduction

Epistemologists have long used intuitions about the epistemic status of agents in hypothetical cases as evidence for theories about knowledge.[1] Indeed the use of intuitions about cases as evidence for epistemological theorizing has had an extraordinarily prominent place in recent debates between linguistic epistemic contextualists[2] (contextualists for short) and invariantists.[3] In particular, contextualists have appealed to intuitions about cases as evidence in offering a theory of the semantics of the verb "know," and in using that theory to help diagnose the problem of Cartesian skepticism.

The root question in the debate is the role that context plays in determining our intuitions about whether an agent has knowledge. Contextualists argue that the standards for knowledge can shift from context to context, depending on what kinds of errors or alternatives are relevant.[4] The evidence for this claim comes from so-called "salience effects." Accordingly, given a situation where an agent intuitively possesses knowledge, merely mentioning an unrealized possibility of error, and thus making it "salient" is enough to weaken or even reverse those intuitions.

Salience effects play an important role in the contextualist's diagnosis of the skeptical problem. Recall a specific form of the argument from ignorance (AI):

1. I don't know I'm not a brain in a vat
2. If I don't know I'm not a brain in a vat, then I don't know I have hands
3. Therefore, I don't know I have hands.

---

[1] Edmund Gettier (1963) is of course the most famous contemporary case. Gettier, as has been widely remarked, does not himself refer to intuitions about cases. The first explicit formulation of the method of using intuitions about cases as evidence for theories comes in Gilbert Harman's (1968: 164) "Knowledge, inference, and explanation."

[2] Epistemic contextualism is a broad term that can be seen as covering many positions. There are two broad camps. Linguistic epistemic contextualists, including Gail Fine, David Lewis, Stewart Cohen, and Keith DeRose base their claims on the semantics of "know." Non-linguistic contextualists, like Michael Williams and David Hume, believe that standards of epistemic appraisal can shift, but that this is a non-linguistic phenomenon. My focus here is linguistic contextualism. Fred Dretske's relevant alternatives theory occupies a kind of middle ground between these two positions (see D. Pritchard, 2002) but is, importantly, defended on the same kind of evidence as linguistic epistemic contextualism.

[3] The key works in the debate being Jason Stanley's (2005) *Knowledge and Practical Interests* (particularly Chapters 1, 2, & 6) and Keith DeRose's (2009) *The Case for Contextualism* (particularly Chapters 1 & 7).

[4] Some contextualists like DeRose, and interest-relative invariantists like Jason Stanley, also believe that the stakes of the action can affect epistemic standards. As stakes don't play a central role in the diagnosis of skepticism, I'll leave the debate aside here. There is a lively empirical debate over stakes effects, see John Turri (*forthcoming*).

Even among the truly deluded, I suspect few genuinely believe they're a brain in a vat, and for the rest of us, I doubt many people think it's even remotely likely.[5] So why think that you don't know you're not a brain in a vat? According to contextualists, the skeptic takes advantage of our general tendency to intuitively discount knowledge ascriptions once an unrealized possibility of error has been mentioned. The force of the skeptical hypotheses about unrolled-out possibilities of error like the Cartesian demon or the alien super-psychologists is that they are global defeaters: ones that cannot possibly be challenged by empirical evidence. Once the defeater is made salient, according to the contextualist, the standards for knowledge shift in the context so that we must rule it out to count as knowing, and because we can't, we have to admit that we do not know we're not a brain in a vat.

Salience effects, then, play a crucial role in the contextualist's diagnosis of the skeptical problem. A genuine diagnosis of the skeptical problem must explain how our ordinary non-skeptical epistemic practices and claims to knowledge are compatible with the intuitiveness of the skeptical argument. Typically we're happy to claim to know many quotidian things, like that we have hands, or that we know where our car is parked. But it also seems, or certainly has been widely claimed, that the argument from ignorance is intuitive - it requires little or no motivation by the skeptic in defense of its premises. The contextualist diagnosis, then, is that in ordinary contexts with ordinary epistemic standards, the threat of global deception is not relevant, and thus it doesn't need to be ruled out in order to know quotidian facts. But, according to this diagnosis, skeptical arguments are based on particularly extreme defeaters, and thus are just extreme versions of salience effects that are much more general. In this way the contextualist offers an account that both explains our ordinary epistemic situation (we know a great deal) and the appeal of AI (it depends on salience effects) by appealing to nothing more than our ordinary epistemic practice.

The solution is a neat one. So-called salience effects have been motivated intuitively by the method of cases: presenting cases where an agent intuitively seems to have knowledge, and pairing it with a further case that mentions a possibility of error, and in which the agent no longer intuitively seems to have knowledge. Keith DeRose's "bank cases," are a classic example of the kind used to intuitively illustrate salience effects. Both

[5] Though, see Nick Bostrom, 2003.

contextualists and invariantists agree that the shift in intuitions exists; where they differ is in their accounts of what determines the shift, and what evidence these intuitions provide for an account for their competing accounts of knowledge. Contextualists follow the ordinary language philosopher's "methodology of the straightforward," which prizes making straightforward claims come out to be true. Thus, these cases become evidence for their position.[6] Invariantists, as we will see later in this chapter, try to argue that the intuitions behind salience cases are byproducts of cognitive biases, and therefore not appropriate evidence for a theory of knowledge.

However, the contextualist diagnosis of skepticism in terms of salience effects faces two challenges. One emerges from recent experimental work into non-philosophers' intuitions about knowledge ascriptions. Contrary to the received view, an early round of experiments seems to show that non-philosophers' intuitions are *not* affected by salience. Even more recent studies seem to show that they are. What should we make of this experimental disagreement? Are knowledge ascriptions sensitive to salience? In sections 2 & 3 of this chapter, I'll take on these questions. Section 2 will look more deeply at traditional presentations of contextualism. Section 3 will take an in depth look at the current experimental debate over salience effects, concluding with a presentation of a series of experiments I've conducted which indicate salience effects are a genuine feature of our ordinary epistemic practice.

The second challenge emerges because contextualism plums commonsense epistemology and our ordinary epistemic practice both for its data and as its foundation for diagnosing the skeptical problem. This diagnosis of skepticism seems to rely heavily on the idea that the salience intuitions of ordinary speakers are not sensitive to the plausibility of error possibilities, even very implausible ones like those mooted by skeptical arguments. However, it is not part of our ordinary epistemic practice to pay much attention to far-fetched contingencies. Indeed, recent work in the psychology of law suggests that it is part of our ordinary practice to positively discount implausible sources of error. This suggests that even if salience intuitions are part of our ordinary linguistic practice, something further is needed to help explain how the skeptic overcomes our *ordinary indifference to implausible error possibilities*. In section 4 I'll lay out this challenge to the contextualist diagnosis. In sections 5

[6] K. Derose, 2009: 153.

and 6 I'll present the results from a series of experiments that show that the salience of an error possibility is a function of its subjective plausibility, and this helps explain the current experimental debate over salience.

In section 7 I will offer an improved diagnosis of the skeptical problem that shows how the skeptic overcomes our ordinary indifference to the implausible. In this section I'll present a series of experiments, which show that this is possible through a sophisticated kind of priming, where error possibilities are presented in sequence from most plausible to least. These experiments, modeled on Descartes' presentation of skeptical error possibilities in *Meditation I*, suggest that priming effects can alter our responses to implausible error possibilities. I will argue that the skeptic overcomes indifference to the implausible through a kind of tacit slippery-slope argument. The result is a richer and empirically legitimate diagnosis of the role of defeaters in skeptical arguments.

2. Linguistic Contextualism

Contextualism is based on two empirical claims. The first is the linguistic claim that the verb "know" is covertly context sensitive. The second is the psychological claim that ordinary speakers exhibit a selective "semantic blindness" to the context-sensitivity of "know," and that this blindness lies at the route of the problem of Cartesian skepticism. Most of the theoretical dispute over contextualism has focused on the linguistic claim, and that will be my focus here. Later, however, I'll explain how the results reported in sections 5 & 6 challenge contextualism's diagnosis of the skeptical problem.

Contextualism comes in many forms. Some contextualists, like Stewart Cohen, claim that as the standards of accuracy in a context vary, so too does the required threshold for justification in that context.[7] Others, like J.L. Austin, David Lewis, and Jonathan Schaffer claim that knowledge is sensitive to the range of contrasts that are relevant in a context.[8] But perhaps the most influential and well-developed view, due to Keith DeRose, is that knowledge depends on the contextually determined range through which one must track the truth.

[7] S. Cohen, 1988, 1999.

[8] J. L. Austin, 1946; D. Lewis, 1996; J. Schaffer 2005a.

In general contextualists claim that "know," a verb, has a semantics similar to scalar absolute adjectives like "flat," "empty," and "tall."[9] For example, the sentence (1) "Tom is tall" might be true in the context of the elementary class where he teaches, while "Tom is tall" might be false in the context of the basketball league where he plays after work. Because "tall" is context-sensitive, the semantic contribution it makes to the sentence that embeds it, and so the truth conditions of the whole sentence, can vary with context.

Contextualists claim that the standards for knowledge can shift from one conversational context to another. So, amidst the low epistemic standards of everyday conversation, it may be true that (2) "Tom knows he has hands," but under the higher epistemic standards of a philosophy seminar, a token of the same sentence might be false.

The evidence for this claim is derived from intuitions about cases. Perhaps the most familiar in the literature are due to Keith DeRose. DeRose argues that our intuitions about hypothetical cases, as it reflects ordinary usage, provides the evidence. He asks his readers to consider the following pair of cases:

Low Standards

> My wife and I are driving home on a Friday afternoon. We plan to stop at the bank on the way home to deposit our paychecks. But as we drive past the bank, we notice that the lines inside are very long, as they often are on Friday afternoons. Although we generally like to deposit our paychecks as soon as possible, it is not especially important in this case that they be deposited right away, so I suggest that we drive straight home and deposit our paychecks on Saturday morning. My wife says, 'Maybe the bank won't be open tomorrow. Lots of banks are closed on Saturdays.' I reply, '**No, I know it'll be open**. I was just there two weeks ago on Saturday. It's open till noon.' (emphasis added)

High Standards

> My wife and I drive past the bank on a Friday afternoon, as in case A, and notice the long lines. I again suggest that we deposit our paychecks on Saturday morning, explaining that I was at the bank on Saturday morning only two weeks ago and discovered that it was open until noon. But in this case we have just written a very large and very important check. If our paychecks are not deposited into our checking account before Monday morning, the important check we wrote will bounce, leaving us in a *very* bad situation. And, of course, the bank is not open on Sunday. My wife reminds me of these facts. She then says, 'Banks do

[9] See P. Unger 1975; S, Cohen, 1999; DeRose, 2009; see also Lynn Murphy (2010: 233) for a discussion of classical examples of scalar absolute adjectives.

> change their hours. Do you know the bank will be open tomorrow?' Remaining as confident as I was before that the bank will be open then, still, I reply, '**Well, no, I don't know.** I'd better go in and make sure.' (emphasis added)[10]

DeRose invites readers to consult their intuitions about whether the emphasized statements (above) are true. In both cases it's stipulated that the bank is open, and in both cases the agent in the vignette has exactly the same information. According to DeRose, and many other epistemologists, it seems clear that in both cases the agent has said something true. DeRose argues contextualism about knowledge is the best explanation of this, seeing as the agent is in no better position to know that the bank is open in the first case than in the second.

Contextualism is an empirical thesis about knowledge ascription. Moreover, it's critical to DeRose's project here, and to contextualism more generally in its goal of diagnosing skepticism, that the evidence for this claim is derived from ordinary epistemic practice. As he says, "The best grounds for accepting contextualism come from how knowledge-attributing (and knowledge-denying) sentences are used in ordinary, non philosophical talk: What ordinary speakers will count as 'knowledge' in some non-philosophical contexts they will deny is such in others."[11] A good diagnosis of the skeptical problem will offer a coherent account of why we seem committed to the things that skeptical arguments deny, and why nevertheless skeptical arguments like AI are so intuitive. Drawing a contrast with high-standards theories of knowledge can help illuminate the situation. If knowledge always required meeting hyperbolic standards, then that would explain the appeal of skepticism, because we don't in fact meet those standards very often, but it wouldn't explain why we ever thought we knew anything in the first place. Likewise, standard fallibilist theories of knowledge why we take a great deal of quotidian knowledge for granted, but they don't explain why skeptical arguments are purported to be so influential.[12] This diagnostic desideratum, in fact, helps explain why investigating the skeptical problem *should* be empirically grounded, and further supports the empirical approach adopted here. We have to know what our ordinary epistemic practices are, what

[10] K. DeRose, 1992: 913.
[11] K. DeRose, 2009: 47.
[12] See M. Williams, 1996: xii - xxiii, introduction.

factors they're sensitive to, and how they vary between groups or personality types, if we're to understand how the skeptic manipulates them.

This pattern of intuitive salience effects is widely accepted in contemporary epistemology, including anti-skeptical invariantists,[13] skeptical invariantists,[14] subject-sensitive invariantists,[15] and of course linguistic contextualists who are in focus here.[16] There is, of course, disagreement between these groups about how best to explain this shifting pattern of intuitions. My concern in this and the following sections is on the pattern of intuitions, keeping it open whether and what mechanisms might generate them. For the time being it is enough to note that while finding evidence of salience effects does not uniquely confirm contextualism, seeing as so many different theories have been designed to explain them, the absence of salience intuitions among ordinary speakers would count as powerful evidence against the contextualist's linguistic hypothesis and their intendant diagnosis of the skeptical problem.

A number of authors have noted that there are two important differences between DeRose's cases, each of which might be driving philosopher's shifting intuitions.[17] The first is the salience of error. In the first case the agent's wife mentions it is possible that the agent's belief is false: "Maybe the bank won't be open tomorrow. Lots of banks are closed on Saturdays." In the second case, the agent's wife explains why, by introducing an unrealized, but uneliminated, possibility of error: "Banks do change their hours. Do you know the bank will be open tomorrow?". The second difference is in the stakes associated with making sure the checks are deposited. In the first case, getting the checks deposited "is not especially important," whereas in the second case failing to do so will leave the subjects "in a *very* bad situation." DeRose has acknowledged this point, and he's endorsed the idea that both stakes *and* salience can change epistemic contexts, though he emphasizes the role of salience, because of its later importance in explaining radical skepticism.[18]

---

[13] P. Rysiew, 2001; J. Hawthorne, 2004; T. Williamson, 2005; A. Hazlet, 2009; and J. Nagel, 2010.

[14] P. Unger, 1971 (though his later work comes closer to contextualism).

[15] J. Stanley, 2005.

[16] In addition to K. DeRose, the list includes F. Dretske, 1970; G. Stine, 1976; D. Lewis, 1979 & 1995; S. Cohen, 1999; J. Shaffer, 2004.

[17] See J. Schaffer, 2004; J. May et al., 2010; J. Schaffer and J. Knobe, 2012.

[18] K. DeRose, 2009: 53 – 59. DeRose does emphasize that the best evidence for contextualism comes from *non-philosophical* contexts, and that contextualism is not, in the first instance, supported by the resolution it offers to the skeptical problem. He says "wise contextualists, however, as we've seen, seek to support their contextualism by appeals to what transpires in ordinary, non-philosophical uses of "know(s)," the most

3.1 Testing Contextualism: Previous Null Results

As contextualism is based on an empirical claim about the semantics of "know," a number of studies have attempted to test the thesis, and the results have been mixed. Three early studies (Buckwalter, 2010; and May, Sinnott-Armstrong, Hull, and Zimmerman, 2010; and Hansen and Chemla, 2013) failed to find a salience effect. These results are balanced by a newer round of results that have detected a salience effect (Schaffer and Knobe, 2012; and Alexander, Gonnerman, and Waterman, forthcoming). Are knowledge ascriptions sensitive to salience? What explains the difference between the early results and the later ones?

In what follows I'll present a detailed summary of previous attempts to test contextualism's semantic thesis. There are two good reasons for doing this. Since the results have been mixed, a detailed look at the differences between the null results and the positive results can illuminate how salience works, and thus how skepticism gets a grip on us. The goal of the comparison will be to identify the factors that influence salience. Defenders of contextualism have criticized the methodology of the experiments that produced null results. I will argue that these criticisms are misplaced. I will further argue that a better explanation of the variation in salience intuitions must track the plausibility of the error possibility.

The second reason for such a detailed comparison is that it permits a detailed look at how people respond to salience. Past studies have focused on mean responses to target questions, and while this is perfectly adequate for assessing the statistical significance of salience manipulations on knowledge attributions, it does not illuminate *how* salience affects knowledge attribution, or whether there are individual differences in intuitions. To understand this, I will argue, we need to focus on modal distribution: doing so shows that salience has a discontinuous "all or nothing" effect, when it occurs at all, and does not operate by undermining confidence. I'll begin by summarizing the past experiments, and then I'll consider and evaluate the methodological complaints leveled against the null results.

MAY, SINNOTT-ARMSTRONG, HULL, AND ZIMMERMAN (2010)

May et al. rightly noticed that DeRose's bank cases involve two variables, stakes and salience, and developed tests that manipulated the variables independently. They presented four vignettes to independent samples: High Stakes – Salient Error, Low Stakes – Salient Error, High Stakes – No Salient Error, and Low Stakes – No Salient Error. In each case,

---

important examples of which are cases that involve no dispute among the parties whatsoever, before *applying* their contextualism to the problem of philosophical skepticism" (DeRose, 2009: 57). He offers this in response to the criticism of Timothy Williamson to the effect that contextualism should be avoided because it offers too easy a resolution of disputes, when some disputes are in fact genuine.

participants were asked the following target question: "How much do you agree with the following statement: 'Hannah knows that the bank will be open on Saturday'." Participants indicated their response on a 7 item likert scale where 1 = strongly disagree and 7 = strongly agree. The stimulus material and response data are reproduced in Table 1 below.[19]

	No Salient Error	Salient Error
High Stakes	Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, and very little in their account, **it is very important that they deposit their paychecks by Saturday**. As they drive past the bank, they notice that the lines inside are very long, as they often are on Friday afternoons. Hannah notes that she was at the bank 2 weeks before on a Saturday morning, and it was open. Hannah says, "I know the bank will be open tomorrow. So we can deposit our paychecks tomorrow morning".	Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, and very little in their account, **it is very important that they deposit their paychecks by Saturday**. As they drive past the bank, they notice that the lines inside are very long, as they often are on Friday afternoons. Hannah notes that she was at the bank 2 weeks before on a Saturday morning, and it was open. **Sarah points out that banks do change their hours.** Hannah says, "I know the bank will be open tomorrow. So we can deposit our paychecks tomorrow morning".
Low Stakes	Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. **It is not important that they do so, as they have no impending bills**. As they drive past the bank, they notice that the lines inside are very long, as they often are on Friday afternoons. Hannah notes that she was at the bank 2 weeks before on a Saturday morning, and it was open. Realizing that it isn't very important that their paychecks are deposited right away, Hannah says, "I know the bank will be open tomorrow. So we can deposit our paychecks tomorrow morning".	Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. **It is not important that they do so, as they have no impending bills.** As they drive past the bank, they notice that the lines inside are very long, as they often are on Friday afternoons. Hannah notes that she was at the bank 2 weeks before on a Saturday morning, and it was open. **Sarah points out that banks do change their hours.** Still, realizing that it isn't very important that their paychecks are deposited right away, Hannah says, "I know the bank will be open tomorrow. So we can deposit our paychecks tomorrow morning".

[19] May et al., 2010: 268 – 270. Emphasis in tables is added.

	No Salient Error	Salient Error
High Stakes	5.07	4.60
Low Stakes	5.33	5.30

Table 1. Stimulus material and main findings from May et al. 2010. Data are mean knowledge attribution to the agent in the vignette, where 1=strongly disagreeing that the agent has knowledge, and 7=strongly agreeing that the agent has knowledge.

May et al. did find a significant effect for stakes: increasing stakes lowered mean knowledge ascriptions. However, and importantly, using the scale's midpoint of 4 as the midpoint for ascribing knowledge, mentioning stakes *did not* move participants to deny knowledge. Moreover, and most importantly for this discussion, not only did salience not push participants to deny knowledge to the agent, May et al. found that salience *does not* even have a significant effect on participant's mean attributions of knowledge. They conclude that these results cast "some doubt on the soundness of arguments for a contextualist or relevant alternatives account of knowledge that appeal to its capturing common-sense or ordinary linguistic practice with respect to *these* bank cases" (emphasis original).[20]

BUCKWALTER (2010)

May et al. use a 2x2 design to compare the interaction of stakes and salience. Buckwalter uses a 3x1 design, comparing high stakes and salient conditions to a neutral condition. In each case, participants were asked the following target question: "On a scale of 1 to 5, circle how much you agree or disagree that Bruno's assertion, 'I know the bank will be open on Saturday' is true." Participants indicated their response on a 5 item likert scale where 1 = strongly disagree and 5 = strongly agree. The stimulus material and response data are in Table 2 below.[21]

[20] May et al., 2010: 272. Unfortunately, May et al. do not report an effect size, or any measures of the distribution of their responses, which makes it difficult to discern the type of effect they generated for stakes.
[21] Buckwalter, 2010: 401 – 404. Emphasis in tables is added.

Control	High Stakes	Salient Error
Sylvie and Bruno are driving home from work on a Friday afternoon. They plan to stop at the bank to deposit their paychecks, but as they drive past the bank they notice that the lines inside are very long. Although they generally like to deposit their paychecks as soon as possible, it is not especially important in this case that they be deposited right away. Bruno tells Sylvie, "I was just here last week and I know that the bank will be open on Saturday." Instead, Bruno suggests that they drive straight home and return to deposit their paychecks on Saturday. When they return to the bank on Saturday, it is open for business.	Sylvie and Bruno are driving home from work on a Friday afternoon. They plan to stop at the bank to deposit their paychecks. Bruno has written a very large check, and **if the money from his pay is not deposited by Monday, it will bounce, leaving Bruno in a very bad situation with his creditors.** As they drive past the bank, they notice that the lines inside are very long. Bruno tells Sylvie, "I was just here last week and I know that the bank will be open on Saturday." Instead, Bruno suggests that they drive straight home and return to deposit their paychecks on Saturday. When they return to the bank on Saturday, it is open for business.	Sylvie and Bruno are driving home from work on a Friday afternoon. They plan to stop at the bank to deposit their paychecks, but as they drive past the bank they notice that the lines inside are very long. Although they generally like to deposit their paychecks as soon as possible, it is not especially important in this case that they be deposited right away. Bruno tells Sylvie, "I was just here last week and I know that the bank will be open on Saturday." Instead, Bruno suggests that they drive straight home and return to deposit their paychecks on Saturday. **Sylvie says, "Banks are typically closed on Saturday. Maybe this bank won't be open tomorrow either. Banks can always change their hours, I remember that this bank used to have different hours."** When they return to the bank on Saturday morning, it is open for business.

Control	High Stakes	Salient Error
3.83	3.71	3.64

Table 2: Stimulus material and results from Buckwalter, 2010. The data represent mean agreement with the claim that a knowledge ascription within the vignette is true: results are on a 5 item likert scale where 1=strongly disagree and 5=strongly agree.

The important feature of Buckwalter's 2010 result, like May et al.'s, is that he does not find a significant difference in the mean responses of participants between the control case and the salient error case. Again, and importantly, using 3 as the midpoint of the scale and construing ratings above it as agreement, all means are significantly above the midpoint, indicating that in general participants *are* as willing to ascribe knowledge even when error possibilities are operative. Buckwalter concludes from his data that, "in the particular bank cases tested we have reason to doubt the contextualist hypothesis; the truth conditions of the knowledge claims tested did not fluctuate between contexts."[22]

[22] Buckwalter, 2010: 403.

The original debate over the contextualist's data developed between philosophers who each offered various hypothetical cases and counter-cases, and not as seen here in empirical studies, of non-philosophers' intuitions. In assessing the implication of these results, it's useful nevertheless to acknowledge a note of caution originally offered by DeRose in the context of professional philosophers trading cases.[23] In such a discussion it's inevitable that some theorists will not share intuitive judgments about a case; DeRose notes that some may not have the contextualist intuition about some cases. But, he argues, to the extent that we do detect context-sensitivity in knowledge attributions, both by appeal to personal intuition (and by experimentation) in at least some cases, the burden is on his opponents to explain those results away. It is not problematic for contextualism that "know" does not exhibit context-sensitivity in all situations: it is context-sensitive so long as it is context-sensitive in some situations.

In the light of this, caution is warranted in interpreting these results. It is very difficult to prove a null result. A single null result is not enough to show that contextualism is false. In the present instance, both Buckwalter, and May et al., are aware of this kind of worry, and are thus modest in their claims. They both recognize that is dangerous to generalize from a single set of results about hypothetical cases, to the conclusion that salience doesn't effect knowledge ascriptions in general. That proviso aside, I think it is notable that they do not find a difference in non-philosophers' willingness to ascribe knowledge with cases that are *extremely similar* to the very cases that have convinced philosophers that the effect does exist. It suggests a new question, too. Do philosophers respond differently to salience than non-philosophers?

3.2 Previous Studies: Significant Results

Perhaps the most immediate question in attempting to interpret these results is just why philosophers have the salience intuition, and non-philosophers don't? The issue is complicated further by more recent results that seem to suggest that non-philosophers *are* sensitive to salience.

[23] K. DeRose, 2009: 177-79.

SCHAFFER & KNOBE (2012)

Jonathan Schaffer and Joshua Knobe defend a moderate kind of contextualism called "contrastivism." According to contrastivism, the traditional formulation of knowledge as a two-place relation between a subject and a proposition such that *S knows P* is false. According to contrastivism the knowledge relation has 3 places, such that a subject knows a proposition relative to a particular contrast proposition such that *S knows P rather than Q*. Because it is a contextualist position that entails a salience effect, according to which making it salient that *Q* is an uneliminated possibility should destroy knowledge, Schaffer and Knobe have to explain away the earlier null results. Their explanation involves two steps. First they posit that when a contrast is not made explicitly salient, speakers use contextual clues to determine the pragmatically implied contrast. They claim that the preceding experiments simply failed to describe a conversational context in which the information given is sufficient to make the error possibility a relevant contrast proposition. They argue,

> The existing studies merely mention the possibility of a change in hours. But merely mentioning a possibility need not render it salient. In particular, the possibility of a bank suddenly changing its hours is a **strange** and **improbable** one, and it might not be possible to make people regard it as relevant merely by mentioning it once in the course of a vignette. Perhaps one can only make people regard a possibility as salient by presenting it in a **concrete** and **vivid** fashion.[24]

Stripped to its bones, Schaffer and Knobe's complaint is that *mentioning* an uneliminated error possibility does not make it salient. To retest the hypothesis, and to remedy what they describe as Buckwalter's "pallid and abstract" presentation of the defeater, Schaffer and Knobe design a vignette that's intended to make "the possibility of the bank changing its hours **concrete** and **vivid**, by using a personal anecdote invested with emotional force."[25] Each conditions was administered independently, and each participant was asked to "rate the degree to which they agreed or disagreed with the attribution: 'Hannah knows that the bank will be open on Saturday." Participants indicated their responses on a seven point likert

[24] Schaffer and Knobe, 2012: 694. Emphasis added.

[25] *Ibid.* Emphasis added. See the discussion in section 4 for an in depth consideration of these features.

scale where 1 = strongly disagree and 7 = strongly agree. The stimulus material and response data are in Table 3 below.[26]

Control	Salient Error
Hannah and Sarah are driving home on a Friday afternoon. They plan to stop at the bank to deposit their paychecks. As they drive past the bank, they notice that the lines inside are very long, as they often are on Friday afternoons. Hannah says, "I was at the bank two weeks before on a Saturday morning, and it was open. So this is a bank that is open on Saturdays. We can just leave now and deposit our paychecks tomorrow morning." Sarah replies, "Ok, that sounds good. Let's go on Saturday."	Hannah and Sarah are driving home on a Friday afternoon. They plan to stop at the bank to deposit their paychecks. As they drive past the bank, they notice that the lines inside are very long, as they often are on Friday afternoons. Hannah says, "I was at the bank two weeks before on a Saturday morning, and it was open. So this is a bank that is open on Saturdays. We can just leave now and deposit our paychecks tomorrow morning." Sarah replies, "Well, banks do change their hours sometimes. **My brother Leon once got into trouble when the bank changed hours on him and closed on Saturday. How frustrating! Just imagine driving here tomorrow and finding the door locked."**

Control	Salient Error
5.54	3.05

Table 3: Schaffer & Knobe, 2012. The data are mean knowledge attribution levels to the agent in the vignette, where 1=strongly disagreeing that the agent has knowledge, and 7=strongly agreeing that the agent has knowledge.

Schaffer and Knobe found a significant difference between the control and defeater condition.[27] Just as importantly, assuming that 3.5 indicates the midpoint of their scale, and that result below can be interpreted as denying knowledge to the agent of the vignette, then their salience condition did push participant's to deny knowledge.[28]

[26] *Ibid*: 694 – 695. Emphasis added in tables.

[27] They do not report an effect size, though the mean difference is large enough to be interpreted as substantial merely by inspection. It is possible to reconstruct one from the data they do report $t(198) = 11.3$; $r = \sqrt{((t^2) / ((t^2) + (df)))}$: $r = .624$, which by Cohen's conventions is a large effect.

[28] Frustratingly, Schaffer and Knobe do not report the standard deviation from the means for this experiment, or any other measure of the distributions. While it is certainly most important that responses in the salience condition do move *on average* below the midline, it would help in diagnosing the effect of salience to understand what kind of distribution of responses they collected. Particularly, the results as reported do not tell us if the effect of mentioning the defeater moved the whole population to the left, essentially shifting a normally distributed pattern of responses below the midpoint, or if it only shifted some responses, creating a bimodal distribution, with some respondents still attributing full knowledge, and others fully denying it. The difference could be of importance in understanding the consequences of salience for the problem of skepticism: if the whole population moves, it suggests that perhaps knowledge is open-textured and comes in degrees; if

HANSEN AND CHEMLA (2013)

There is an open and larger debate in the philosophy of language over the exact nature of the contribution that context makes to meaning.[29] Nat Hansen and Emmanuel Chemla worry that the methods employed in previous tests of epistemic contextualism are flawed both as tests of the epistemic doctrine and the wider linguistic one. Their first complaint, which is well taken, is that the null results of Buckwalter and May et al. have been interpreted as evidence *against* contextualism, when no such conclusion is warranted by the failure to find an effect. The basics of experimental design and statistical analysis *assume* that there are no differences between conditions, so the failure to find an effect for a manipulation does not show that a particular manipulation *cannot* have an effect. Experimental tests are tests of relationships between variables, and there can be many reasons why one particular test does not detect an actual relationship.[30] A good hypothesis is that heating one end of a thermometer causes mercury to rise. Heating one end of a thermometer while the other end is immersed in ice, and finding the mercury doesn't rise, doesn't disconfirm the hypothesis.

Hansen and Chemla's point is well taken, and it is true that in general null results are interpreted very cautiously, and are generally interpreted as inconclusive within the social sciences.[31] They are, however, a little unfair to Buckwalter, who recognizes this and interprets his results very cautiously. What is fascinating and interesting about Buckwalter's null result is that it uses *scenarios that are very similar* to the ones that otherwise generated wide consensus among philosophers. Now, a study that focuses only on philosophers' intuitions about bank cases has not been done to my knowledge, but if we assume that the widely *discussed* professional consensus exists in fact, then we would have the basis for a meaningful

---

individuals move from one pole to the other, it suggests that knowledge is a threshold concept, and salience is an all or nothing affair, and there is merely individual variation in where that threshold is set.

[29] See C. Travis (1997), S. Gross (2001), F. Recanati, (2004), and H. Cappelen & E. Lepore (2005).

[30] Hansen and Chemla, 2013: 292 – 293.

[31] This is perhaps an appropriate place to address a pervasive perception of experimental philosophy as merely negative. It is often said that x-phi is sterile and unproductive because it merely aims to disprove the work of past theorists. This is, perhaps, true of some work, but by in large the experimental methods of the social and psychological sciences do not really provide philosophers with the tools to do any such thing. It is very, very difficult to defend a null result, which is what sterile disproof requires. Rather, the most negative strand of x-phi, the restrictionist challenge, is premised on the positive result that there are significant differences in individuals responses to classic thought experiments that can be predicted by variables that have classically been assumed to be irrelevant, namely race, socio-economic status, culture, and ordering. X-phi very rarely tries to show that the *intuitions aren't there*, rather it tries to show that unexpected ones *are there*. Whether that is positive or negative just depends on what theory you advocated previously.

discussion of a positive result, namely the difference between philosophers and non-philosophers.[32]

Keith DeRose has objected that the designs of Buckwalter and May et al.'s are not a fair test of contextualism.[33] I'll discuss his objections in more detail below, but here they are useful for understanding Hansen and Chemla's experiment. The bank cases that DeRose presents throughout his work on contextualism feature pairs of cases, LOW and HIGH, where in the former the agent in the vignette claims "I know the bank will be open," but in the latter claims "I don't know the bank will be open." According to DeRose the best explanation of the intuition that *both* sentences are true, as uttered by the subject, is contextualism. In the studies of May et al., and Schaffer and Knobe, participants are asked if the subject in the vignette has knowledge, rather than if one of their utterances is true. According to DeRose this is not a fair test for two reasons. The first is that by asking about knowledge, the experiment doesn't specify the context of evaluation (whether it is the subject's or the participant's) whereas focusing on an utterance does fix the context as the context of the speaker. Without clearly fixing the context, according to DeRose, it is possible that the attributor's context is more or less demanding than the context that is being evaluated. Second, DeRose points out that because of the phenomenon of conversational accommodation, discussed by David Lewis,[34] where speakers are generally inclined to charitably interpret a speaker's utterance such that it has content which makes it true, we should favor designs where the speaker says something true. Thus, the contextualist prediction, leaving aside accommodation, in Buckwalter's design is to say that the subject said something true in the LOW case but something false in the HIGH case. But DeRose worries that because accommodation is such a powerful force, and can in fact act to determine what the conversational context is, by having the subject of the vignette say "I

[32] Such a study of philosophers' intuitions of bank cases has not been done. However, David Bourget and David Chalmers (forthcoming) descriptive study of philosophers beliefs provides some foundation for this claim. They found that a plurality of philosophers (both specialist and non-specialists in epistemology) described themselves as contextualists about knowledge (40%) when choosing between it, invariantism (30%), or other (30%) (pg. 12 mnscpt.). In so far as it's plausible to assume that contextualists have contextualist intuitions about bank cases, and that at least some invariantists do to, then there is at least some evidence that bank case intuitions are widely shared among philosophers. This makes their absence among non-philosophers interesting, and not merely a null result, but a real difference that raises a question.

[33] K. DeRose, 2011: 83 - 88.

[34] D. Lewis, 1979: 346 - 7.

know the bank will be open" the context is actually converted by the utterance into a LOW case.[35] This then would not be a fair test of contextualism.

Hansen and Chemla do focus their studies on the truth of statements. However, they note in conjunction with their earlier worries about Buckwalter's interpretation of his result, that DeRose's proposed study design aims to confirm contextualism through a null result, i.e. no difference between LOW affirmation of knowledge and HIGH denial of knowledge. But such a result could be produced by many factors, and not just the interaction between context and "know." Even at the most superficial level the two conditions have two differences, context and the use of negation, and either could plausibly explain the lack of difference DeRose predicts.

To remedy this Hansen and Chemla prepare four kinds of cases to test epistemic contextualism. They prepare two versions of LOW and two versions of HIGH each with a subject's assertion expressed in terms of what they call "positive polarity" for affirming knowledge and "negative polarity" for denying knowledge (see Table 5 below).[36] They also prepared 3 other vignettes to test the generality of the epistemic contextualist's predictions. They further prepared 4 tests involving color attribution drawn from Charles Travis' work, and 2 comprehension checks. Given their design, Buckwalter's presentation would conform to the left column, while DeRose's preferred design would be on the diagonal. Participants were then asked to assess the truth value of the subject's indicated statement on a continuous 100 point scale anchored on the left with "false" and on the right with "true." Strict invariantism predicts there should be no difference between any of these conditions.[37] Strict contextualism predicts that participants will tend to find LOW – POSITIVE and HIGH – NEGATIVE truer (on the scale) than LOW – NEGATIVE and HIGH – POSITIVE. These predictions, assume that the context is interpreted as it's coded, and they also assume that accommodation is not at work in determining the standards of the conversational context.

Including the affects of accommodation is more difficult. Pure contextualism predicts that you say something false in LOW – NEGATIVE because there is no reason to deny knowledge in such a simple case. Likewise, it predicts that in HIGH – POSITIVE claiming knowledge is also false, because the standards have gone up. Predicting the results

[35] K. DeRose, 2011: 88.

[36] N. Hansen and E. Chemla, 2013: 299. Emphasis in charts added to indicate differences between conditions.

[37] Typically, of course, invariantists allow that salience intuitions occur, and thus must be explained as an error.

on the assumption that contextualism and accommodation are active is more difficult because one needs to first specify the relative power of these forces in the given context. The HIGH – POSITIVE case brings the conflict to light. This is a key case for the contextualist because it helps explain the appeal of radical skepticism: when standards are at Cartesian levels, we agree with the skeptic that our knowledge claims are false, because we cannot meet their demands by ruling out the skeptical hypothesis. Likewise in these vignettes, the subject cannot rule out the defeater that has been introduced. Now, pure contextualism, on the assumption that introducing the error possibility raises the contextually determined standard, predicts that we should judge Bruno's statement to be false (see below). But, according to accommodation, we try to charitably interpret the content of the utterance in such a way as to read it as true. In this way, Bruno's statement tends to lower the contextual standards back to the point at which his statement is true. As these two forces oppose each other, and this is one of DeRose's complaints about the presentation in Buckwalter, the prediction we make will depend entirely on how powerful we believe each of these forces is in the given vignette. Below I've followed DeRose's lead in assuming that "accommodation [is] a very important force in the interpretation of context-sensitive language."[38] Table 4, below, suggests theoretical predictions for responses to Hansen and Chemla's vignettes.

Theoretical Predictions for Conditions	Stakes	Positive	Negative
Invariantism	Low	Same	Same
	High	Same	Same
Contextualism	Low	True	False
	High	False	True
Accommodation	Low	True	True
	High	True	True
Results[39]	Low	True (~73%)	True (~69%)
	High	True (~65%)	True (~74%)

Table 4: Theoretical predictions for various epistemological theories, based on the design of Hansen and Chemla (2013), and results.

[38] K. DeRose, 2011: 88. The relative power of accommodation and salience effects is, I will argue, a serious empirical question that needs to be addressed to offer an adequate diagnosis of the skeptical problem. A genuine diagnosis of the skeptical problem will explain why salience effects are *not* cancelled by accommodation in skeptical contexts when they are, according to DeRose's critique of tests of contextualism, in contexts with non-skeptical hypotheses.

[39] N. Hansen and E. Chemla, 2013: 304. The numbers I report are estimates, using visual measurement tools, from graphical reporting in their paper. They do not numerically report mean response for their experiments. While I'm unsure about the unit estimates, I'm very confident that all cases are above the midline.

	Positive Polarity	Negative Polarity
Low Stakes Low Salience	Slyvie and Bruno are driving home from work on a Friday afternoon. They plan to stop at the bank to deposit their paychecks, but as they drive past the bank they notice that the lines inside are very long. **Although they generally like to deposit their paychecks as soon as possible, it is not especially important in this case that they be deposited right away.** Bruno suggests that they drive straight home and return to deposit their paychecks on Saturday morning. He remembers driving by last Saturday and seeing that it was open till noon. **Slyvie says, "Maybe the bank won't be open tomorrow. Lots of banks are closed on Saturdays. On the other hand, shops are often open on Saturdays in this neighborhood. Do you know the bank will be open tomorrow?"** **Bruno replies, "I know the bank will be open tomorrow."** It turns out that the bank is open on Saturday	Slyvie and Bruno are driving home from work on a Friday afternoon. They plan to stop at the bank to deposit their paychecks, but as they drive past the bank they notice that the lines inside are very long. **Although they generally like to deposit their paychecks as soon as possible, it is not especially important in this case that they be deposited right away.** Bruno suggests that they drive straight home and return to deposit their paychecks on Saturday morning. He remembers driving by last Saturday and seeing that it was open till noon. **Slyvie says, "Maybe the bank won't be open tomorrow. Lots of banks are closed on Saturdays. On the other hand, shops are often open on Saturdays in this neighborhood. Do you know the bank will be open tomorrow?"** **Bruno replies, "Well, no, I don't know the bank will be open tomorrow. I'd better go in and make sure."** It turns out that the bank is open on Saturday
High Stakes High Salience	Slyvie and Bruno are driving home from work on a Friday afternoon. They plan to stop at the bank to deposit their paychecks, but as they drive past the bank they notice that the lines inside are very long. **Bruno and Slyie have just written a very large check, and if the money from their pay is not deposited by Monday, it will bounce, leaving them in a very bad situation with their creditors. And, of course, the bank is not open on Sunday.** Bruno suggests that they drive straight home and return to deposit their paychecks on Saturday morning. He remembers driving by last Saturday and seeing that it was open till noon. **Slyvie reminds Bruno of how important it is to deposit the check before Monday and says, "Banks are typically closed on Saturday. Maybe this bank won't be open tomorrow either. Banks can always change their hours, I remember that this bank used to have different hours. Do you know the bank will be open tomorrow?"** **Bruno replies, "I know the bank will be open tomorrow."** It turns out that the bank is open on Saturday.	Slyvie and Bruno are driving home from work on a Friday afternoon. They plan to stop at the bank to deposit their paychecks, but as they drive past the bank they notice that the lines inside are very long. **Bruno and Slyie have just written a very large check, and if the money from their pay is not deposited by Monday, it will bounce, leaving them in a very bad situation with their creditors. And, of course, the bank is not open on Sunday.** Bruno suggests that they drive straight home and return to deposit their paychecks on Saturday morning. He remembers driving by last Saturday and seeing that it was open till noon. **Slyvie reminds Bruno of how important it is to deposit the check before Monday and says, "Banks are typically closed on Saturday. Maybe this bank won't be open tomorrow either. Banks can always change their hours, I remember that this bank used to have different hours. Do you know the bank will be open tomorrow?"** **Bruno replies, "Well, no, I don't know the bank will be open tomorrow. I'd better go in and make sure."** It turns out that the bank is open on Saturday

Table 5: Stimulus material from Hansen and Chemla, 2013.

Hansen and Chemla's results are a mixed bag for epistemic contextualism. They do find significant contextual variation in their non-epistemic vignettes, supporting more general semantic contextualism. These effects, however, are much larger than the ones they were able to produce with their epistemic vignettes.[40] Importantly, Hansen and Chemla did find a significant difference between contexts, controlling for the polarity of the response, and this result is generally in line with the predictions of the contextualist. However, Hansen and Chemla introduce an important caveat in interpreting their results. As mentioned, they present four different vignettes of the epistemic schema in random order to see if contextualist intuitions generalize. This means that each participant saw four different vignettes and all four possible combinations of context and polarity in random order. Consequently, only most of their judgments were made having seen a "bank" case before. Hansen and Chemla found that when they focused their analysis only responses to the first appearance of epistemic cases, they found that the significant result disappeared, once participants no longer had cases to contrast with it. Thus, when no contrast was possible, and incorporating DeRose's suggestion that both *stakes* and *salience* are part of the manipulation between LOW and HIGH, and using his preferred presentation focusing on truth assessments rather than knowledge ascriptions, Hansen and Chemla reproduce Buckwalter's null result.[41]

In summing up their results, they look very much like they support the "accommodation" prediction in table 4. But as accommodation is indistinguishable from a null result, it doesn't offer any support to the contextualist. More importantly, they do not report any measures of the distribution of responses, obscuring whether or not the participants evaluations were normally distributed or not. Given that they use ANOVA to assess the difference in their means, we can cautiously assume that their data were normally distributed, as this is an assumption of the *F* statistic. If this is right, it further damages the contextualist's case. Assuming that 50% on Hansen and Chemla's scale was interpreted as

[40] Although Hansen and Chemla report that "the contextual effect is weaker for the knowledge scenarios than for the color and miscellaneous scenarios," they do not report effect size measures that would give some sense of the size of the difference, nor can the appropriate measure of effect size here, n^2_p, be recovered just the reported *F* statistic and *df*. While their design is generally much superior to any work on epistemic contextualism that I'm aware of, their data reporting is somewhat wanting. They visually report mean responses for their knowledge scenarios with bar graphs. This makes it very tough to even get a sense of the mean responses in their four types of context and polarity, but visually inspecting the chart on pg. 304 suggests that overall responses were between 60 and 70% on a 100 unit scale.

[41] N. Hansen and E. Chemla, 2013: 306 – 307.

by their participants as the boundary between assessing the subjects' speech acts as true or false, then not only are the means *not* statistically different when encountered for the first time, all the knowledge cases remain above the midline. In their non-epistemic vignettes, there was a statistically significant difference between contexts, and responses crossed the midline even though contrasts were not available.[42] This suggests that even when it is possible to make contrasts between cases with epistemic vignettes, it is not enough to generate differences in knowledge attribution.

There are a few general lessons from their study. First, their results suggest that perhaps salience intuitions depend on contrasts; following their own advice that null results should be interpreted carefully, we can say that intuitions generally supportive of epistemic contextualism were only detected when a *contrast* between contextual conditions was possible. Second, even in cases where contrasts were possible, the variation in responses did not cross the midline between contexts, suggesting context can affect confidence, but is not sufficient to affect the denial or attribution of knowledge. And third, as Hansen and Chemla's contextual high-standard conditions used *both* stakes and salience, it is not possible to determine if either feature in isolation was responsible for the contrast effects. While this indeterminacy is not especially meaningful to the case for generic epistemic contextualism, it is especially relevant if one aims to apply contextualism to a diagnosis of radical skepticism, where matters of stakes are not in play. It is often remarked that global skeptical scenarios do not have practical consequences, for if we accept the skeptic's case, and we cannot tell between the skeptic's hypotheses and the real world hypotheses, there's no reason to live your life differently. It's perhaps open to the contextualist to contest this, but then their diagnosis of the skeptical problem would run through commitments that are more contentious than their present reliance on ordinary epistemic practice. And third, even in cases where contrasts were possible, the variation in responses did not cross the midline between contexts, suggesting context can affect confidence, but is not sufficient to affect the denial or attribution of knowledge.

[42] N. Hansen & E. Chemla, 2013: 304.

ALEXANDER, GONNERMAN, & WATERMAN (*FORTHCOMING*)

Jennifer Nagel has proposed that salience effects are in fact the result of a psychological bias known as epistemic egocentrism (aka "the curse of knowledge") and not context-sensitivity. Epistemic egocentrism refers to a pervasive tendency to project our beliefs and other mental states onto others. Nagel argues that because of epistemic egocentrism, when an unrealized possibility of error is introduced about an agent's belief, we treat our concerns as theirs, and we penalize them for failing to respond to those concerns appropriately. Nagel's proposal is attractive to invariantists because it offers a way to explain salience effects as the *product* of a bias.[43]

In forthcoming paper, Josh Alexander, Chad Gonnerman, and John Waterman, set out to test Nagel's hypothesis using a number of variations on vignettes originally introduced by Stewart Cohen.[44] Their results are of particular interest to the discussion of skepticism because they are the first tests of pure salience effects in non-bank cases.[45] Participants in their study received one of two vignettes, and were then asked: "indicate the extent to which you agree or disagree with the claim that 'John knows that the table is red'." Answers were assessed using a six-point Likert scale with 1 = strongly disagree and 6 = strongly agree. Their stimulus material and response data are collected in Table 6.

Control	Salient Error
John A. Doe is in a furniture store. He is looking at a bright red table under normal lighting conditions. He believes the table is red.	John B. Doe is in a furniture store. He is looking at a bright red table under normal lighting conditions. He believes the table is red. **However, a white table under red lighting conditions would look *exactly* the same to him, and he has not checked whether the lighting is normal, or whether there might be a red spotlight shining on the table.**

Control	Salient Error
5.5	3.78

Table 6: Alexander at al. *forthcoming*: The control and Salient Error cases were presented independently. Data are mean agreement ratings with the statement "John knows the table is red," where 1=strongly disagree and 6=strongly agree

[43] J. Nagel, 2010: 303.

[44] S. Cohen, 2002.

[45] Buckwalter (2010) looks at salience in banks, May et al. (2010) look at salience and stakes independently in banks, Feltz and Zarpantine (2010) look at salience and stakes together in banks and non-banks, Schaffer and Knobe look at salience in banks, but confound the result by adding stakes; Hansen and Chemla look at salience and stakes together in banks and non-banks.

Consistent with Schaffer and Knobe's results, they found that participants were significantly more willing to attribute knowledge in the control ($M = 5.50$, $SD = 1.14$) than in the salient error condition ($M = 3.78$, $SD = 1.40$). Even with a small sample, the difference between the conditions was significant ($t(38) = 4.29$, $p < .001$), and the effect size (Cohen's $d = 1.36$) was large by Cohen's guidelines. Numerous variations on the same basic vignettes with larger samples confirm the salience result and the relative size of the effect.

A few important differences between Alexander et al.'s results and previous work, especially Shaffer and Knobe's, deserve mention. First, introducing a new scenario is important. While it certainly makes sense to test the contextualist's own cases, and variations thereof, any conclusions that could be validly drawn from them is severely limited. It is certainly conceivable that salience judgments are at least partially sensitive to background knowledge. If it just is implausible to most non-philosophers[46] that banks change their hours frequently, and if salience judgments are at least somewhat responsive to plausibility, then these cases would be unlikely to detect any effect.

A second difference is that Alexander et al.'s salience vignette only presents a possibility of error, while Schaffer and Knobe's reintroduces a practical or emotional element to the evaluation. Although Schaffer and Knobe were trying to test pure salience effects, the only way in which their study differs from Buckwalter, and May et al.'s is in introducing *stakes* to the evaluation.[47] Buckwalter, May et al., and Schaffer and Knobe all mention that "banks do change their hours." But where the previous studies stop there, the way in which Schaffer and Knobe attempt to improve their vignettes by presenting them in a "concrete and vivid fashion,"[48] is to add a description of how one of the subject's brothers "got in trouble" when the bank changed its hours and that this was "frustrating."[49] Schaffer and Knobe's aim seems to be to make the unrealized defeater "concrete" by presenting an instance in which that defeater is realized. Making it concrete in this way may well have been what drove the salience effect. But further study would be needed to distinguish whether

[46] Or, perhaps, less familiar to *college students*. I will note, just in passing, that all the negative salience results were conducted on college undergraduates, while the two studies supporting it were conducted on adults (outside an Australian mall in Schaffer & Knobe's case, or via Amazon Mechanical Turk in Alexander et al.'s case). It may just be that undergraduates are neither familiar nor invested in the hours of banks. Josh Alexander (2012: 45) notes that one of the central results of experimental philosophy is that intuitions can vary by culture, so it is not an insignificant detail that Schaffer an Knobe's subjects were all Australian.

[47] J. Alexander makes this point clearly, 2012: 46.

[48] Schaffer and Knobe, 2012: 694.

[49] *Ibid,* 694 – 695.

this was the relevant difference, or the practical consequences for the subject's brother, or the invitation to emotionally engage with the vignette. Neither of these later two factors is epistemic, but they are meaningful differences between the two experimental conditions. Moreover, the hypothesis that these two differences, and thus stakes, drove their result is just as plausible, given that May et al. did find a small effect for stakes. Alexander et al.'s vignettes, on the other hand, do demonstrate a clear salience affect by avoiding any mention of stakes in introducing the unrealized possibility of error.

This same point holds when comparing Alexander et al.'s vignettes to Hansen and Chemla's. As Hansen and Chemla acknowledge, their experiments cannot resolve the relative contributions of stakes and salience in driving the context effect that they found. More importantly, and unlike Hansen and Chemla's, Alexander et al.'s results do not depend on any potential contrast-effect in intuitions, because each participant in their study received only one case.

A third important difference is that while Alexander et al.'s studies produce effect-sizes that are statistically comparable to Schaffer and Knobe's, their table and light studies do not push *mean* responses below the theoretical midpoint. This could be interpreted in one of two ways. Focusing solely on mean responses, the natural inference is to interpret this result as supporting the idea that salience *can* affect confidence. These results would not, however, support the claim that context actually affects knowledge attribution, because the means don't go below the midline.

This would be an appropriate inference if the responses were normally distributed, like a bell curve. When responses are normally distributed, the mean is an indication of the central tendency of participants' responses, and thus, *ceteris paribus,* the effect of salience on participants. Alexander et al. helpfully report the skew and kurtosis levels of their samples, which show that responses to their vignettes are not normally distributed.[50] The control case is heavily left skewed, while their salience case is bimodal around the midpoint of the scale – the major mode is above the mid-line and the minor mode is below it, explaining why the mean is slightly above 3.5. This suggests that salience affects knowledge attribution in a way that is more complicated than a linear decrease in confidence. If salience did affect confidence in a linear way, you would expect to see the curve move as a whole, down the

[50] Alexander et al. footnote 8, pg. 28 mnscpt.

Likert scale. Knowledge would be affirmed when the modal response is above the midline, and denied when it drops below. Individual variation explains the left and right tails of responses.

The clustering around two modes (see below, Figure 1.) observed by Alexander et al. suggests that the effect of salience on knowledge attributions is discontinuous. This is a small sample, however. A meta-analysis[51] (see below, Figure 2.) of Alexander et al.'s results and my own studies that included Nagel's control (the simple vignette) and salient error cases (the elaborate vignette) shows that the bimodal distribution remains. There is no definitive test for bimodality,[52] and the general tests for violating normality cannot be applied to large samples, such as this one,[53] so a perfectly acceptable test is still visual inspection. What should we make of this? The best explanation is that salience does not act on knowledge attribution merely by lowering confidence, but that participants tend to either ascribe knowledge or not, depending on whether or not they take the defeater to be salient. If they judge that it's salient, they do not ascribe knowledge to the subject; if they do not judge that it's salient, they do ascribe knowledge to the subject. This is a point I'll return to in the next section. Composition matters: the mode is the message.

One might worry that salience effects in studies like these are just a matter of accommodation. The experimental materials mention something, and the experimental subjects feel the compulsion to respond to it. Indeed, accommodation may be just what is going on in normal conversations - a speaker mentions a possibility and we feel compelled to take it into account! But then that is a feature of the phenomenon, not a bug! The worry, then, is that the experimental condition prompts accommodation where we would not find it in ordinary conversation. Accommodation is always a worry in any empirical study, and it cannot be ruled out *a priori*. As we'll see in the experiments below, and more general results

[51] These results are taken from four studies, each of which used Nagel's simple and elaborate vignettes as controls. A meta-analysis of the results confirms Alexander et al.'s findings; knowledge attributions using the simple vignette as the control (n=134, M=4.97, SD=1.14) were significantly greater ($t(280)$= 9.36, p< 0.001) than in those using the elaborate vignette as the salience condition (n= 148, M = 3.46, SD = 1.52). In this case the mean *is* slightly below the midline, but not in a meaningful way. Because the skewness of the elaborate vignette's distribution=-0.13, and its kurtosi =1.82, a t-test is not entirely appropriate. A more appropriate non-parametric analysis using the Mann-Whitney test (U = 15560, p < 0.0001) confirms the result.

[52] Frankland, B. W. & B. Zumbo, 2002: 166, "Quantifying Bimodality Part I: An Easily Implemented Method Using *SPSS*," *Journal of Modern Applied Statistical Methods*, Vol. 1, No. 1: 157 – 166.

[53] A. Field, 2009: 139.

discussed in Chapter 5, checking the salience pattern against more general response patterns does not indicate accommodation in response to these questions.

Summing up, there are a number of important lessons that can be learned from Alexander et al.'s results. Their results support the conclusion that salience effects don't depend on contrasts, or on stakes. Their results support the conclusion that salience can affect knowledge attributions, not merely confidence, because their response distributions are bimodal across the midline of their scale.[54] And finally, by varying the content of their vignettes their results support the conclusion that *background knowledge*, and not the mere presence of an error possibility, plays a role in salience effects.

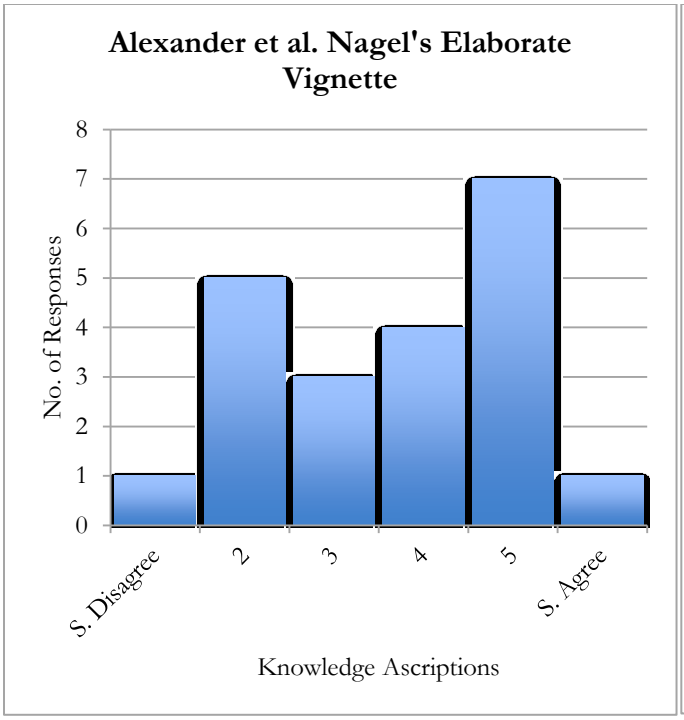


Figure 1.

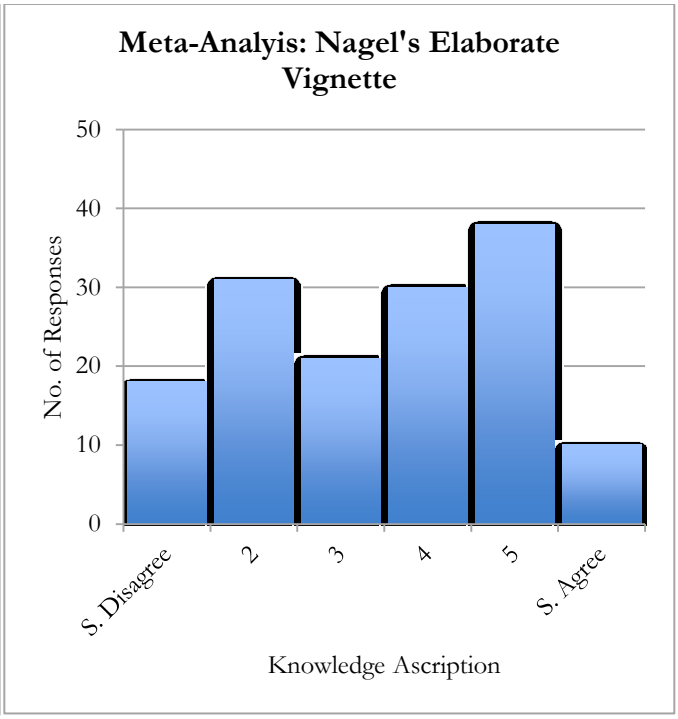


Fig. 2.

3.3 A Puzzle: Bad Methodology?

How do we explain the difference between the early null results and the later positive results? There are a couple of possibilities. The methods that produced the null results have been vigorously critiqued by DeRose; his criticisms are a natural place to look first for an explanation of why the early tests failed to find a salience effect. I'll argue that these worries

[54] Another way to put this is that "knowledge is all or nothing": salience does not diminish confidence in a smooth, linear fashion, but it either causes the denial of knowledge or has little effect.

are misplaced: they do not point to serious methodological flaws, and they do not explain the difference between the two sets of results. Moreover, I'll explain why DeRose should hope his criticisms *are not* the explanation of the differences, because if he's right, then the basic phenomenon that gives rise to the skeptical threat is a very fragile one, and this kind of instability is at odds with how influential skeptical arguments have historically been.

After this I'll argue that the difference is better explained by the actual salience of the defeaters presented to participants, and not merely idiosyncrasies of the methods employed in presenting them. More specifically, I'll introduce the idea that salience is at least partially a function of participants' background knowledge and their subjective assessment of the probability of the potential error possibility. I'll defend this view experimentally in sections 5 and 6.

Perhaps the most immediate explanation of the difference between the null results and the positive results is the one offered by Hansen and Chemla: null results aren't really *results*. They aren't evidence against contextualism; they just aren't evidence for it. As I said before, this is important to acknowledge, but it shouldn't silence our attempts to understand the factors that go into salience effects. It is important to question *why* the philosophical consensus about bank cases is so different from the non-philosophers'.

Keith DeRose critiques the null results on a number of fronts. The lesser of his worries is the possibility that the null results are a product of poor survey design, and he cites Simon Cullen's 2010 critique of the methodology used by experimental philosophers defending the restrictionist challenge as support. Restrictionism claims that the experimentally demonstrated difference between folk intuitions and philosophers' intuitions about various high-profile thought experiments, like Gettier cases, undermines the credibility of appeal to intuitions as evidence. Cullen argues that the results restrictionists appeal to are the product of poor survey methods, rather than genuine differences in intuitions about important thought experiments.

To make his case he points to two areas in which poor design can result in just the pattern of responses restrictionists have found. The first point begins by emphasizing the importance of pragmatic factors in accurate interpretation, and then draws attention to all the ways that paper and internet surveys are strange and impoverished conversational contexts when compared to actual conversations. As a result, without proper instruction and precaution, surveys can lead to misunderstandings that invalidate the data they generate.

A possible example is that without the philosopher's background focus on the nature of knowledge, a naïve participant in a survey might mistakenly interpret one of DeRose's bank cases as an inquiry about *how often* banks change their hours, and not about the state of knowledge of the subject in the vignette. Confusion instilled by impoverished contextual clues can cause participants to answer at random, which almost perforce will differ from the response patterns of professional philosophers

Cullen's second point is that choices about response type (interval vs. binary data) can determine very different patterns of responses. This second worry we can set aside, as all the studies here use interval measures. The first consideration, though, is particularly relevant just because epistemic contextualism claims that features of the conversational context *fix* the standards of evaluation. A fair test of contextualism demands that this aspect of survey administration is considered closely.

DeRose's more serious critique of the null results is that the surveys are presented in a way that does not actually test contextualism's predictions. This concern slots directly into Cullen's worry that surveys are impoverished conversational contexts. In particular, he argues the null results are flawed because of they fail to focus on i) the truth values of agent's assertions, ii) that they fail to pair cases as affirmations of knowledge in the control condition with denials of knowledge in the salience condition, and iii) they do not ask about stakes and salience together.[55] I'll present DeRose's worries briefly below. Then, I'll explain that none of the worries he raises actually explains the difference between null results and the positive results. I'll conclude this section by arguing that DeRose's position that the only fair test of contextualism satisfies the above three conditions threatens to undermine the diagnosis it offers of skepticism by reducing contextualism to a quirky feature of usage, rather than a robust and serious proposal about how knowledge ascription works.

DeRose's first concern focuses on truth-values. He argues that asking participants whether or not the subject in the vignette has said something true when they make a knowledge affirming or denying utterance is preferable to asking the participants if the subject knows or not. DeRose is an *attributor* contextualist, which is to say that it is a consequence of his view that two people in different conversational contexts can appropriately assign different truth-values to the same knowledge affirming utterance made

[55] He is so serious in this critique that he claims, "I don't think as [Schaffer and Knobe] do that any clear predictions of contextualism are vindicated by their survey." K. DeRose, 2011: 92.

by a speaker, so long as the attributors respective contexts put in place different epistemic standards. As such, when the evaluator and the speaker being evaluated are in different contexts, it is sometimes appropriate for the evaluator to use their own context, while in others it is appropriate to use the speakers. Thus, according to DeRose, using surveys to assess contextualism is valid on the assumption that each of the participants is in a context with roughly the same standards, but nothing about merely *reading* the same vignette guarantees this. If participant's contexts vary unpredictably, then a null result is the expected outcome.

To avoid this, DeRose argues, it is essential to focus the participant's task on assessing the truth values of the agent's utterances in the vignette, because this helps focus attention on the agent's context as the appropriate context for evaluation.[56] This is accomplished through the phenomenon of conversational accommodation, discussed by David Lewis,[57] where speakers are generally inclined to charitably interpret a speaker's utterance such that it has content that makes it true. This will tend to put the focus on the speaker's context of evaluation. It's worth noting that this is an empirical claim, but a plausible one.

DeRose's second complaint about the null results is related to the first. He has consistently presented his cases such that in LOW the subject affirms knowledge, "I know the bank will be open on Saturday," while in HIGH the subject denies knowledge, "I don't know that the bank will be open on Saturday." The asymmetry is by design: without it the assessment of the utterance in HIGH wouldn't benefit from the force of accommodation.

Finally, DeRose contends that the best way to test contextualism is to test its best cases, and these are ones where *stakes* and *salience* are acting together to raise the standards of HIGH. While he acknowledges that such a design cannot distinguish the contributions of

---

[56] In his 2009 DeRose adverts to Jennifer Nagel's (2007) discussion of an emerging doubt among epistemologists about the coherence of epistemic intuitions about knowledge attribution. Nagel's discussion uses the messy and inconclusive nature of the Gettier debate to suggest that intuitions about knowledge attributions might not yield a consistent analysis of knowledge. DeRose also points to his own weak intuitions about the fake barn cases, and Ruth Millikan's critique of them, in suggesting why contextualist cases should focus on evaluations of knowledge affirming sentences, rather than attributions of knowledge. But if the incoherent nature of our concept of knowledge is the problem, it's not obvious why we should prefer evaluations to attributions, absent any other forces (but see the discussion of accommodation in the body). If "know" is an incoherent mess, and explodes in use when hypothetical scenarios get complicated, why would focusing on the sentence level fix that?

[57] D. Lewis, 1979: 346 - 7.

these two variables in changing the contextual standards, DeRose believes the first concern should be testing for the effect, not how it arises.[58]

DeRose is certainly right, the various experimental tests of contextualism do not conform to his preferred design, but can we explain the difference between the null and positive results in terms of them? I think it is unlikely. Table 7, below, summarizes how each of the experiments I've discussed fairs according to the design favored by DeRose. The three columns correspond to the important design features DeRose discusses. If an experiment conforms to the preference, it gets a check. It turns out that salience effects don't turn on any of the factors DeRose highlights. The important contrast that brings this to light can be seen in the last two rows. Hansen and Chemla follow DeRose's preferred design, and they do not find context effects when they constrain their analysis initial responses. Alexander et al., however, do not follow the recommendations and yet they do find a salience effect.

Experiment	Evaluation	True - False Asymmetry	Stakes and Salience
May et al.	✗	✗	✗
Buckwalter	✓	✗	✗
Schaffer & Knobe	✗	✗	✓†
Hansen & Chemla	✓	✓	✓
Alexander et al.	✗	✗	✗

Table 7: DeRose's preferred design features

This is not to say that DeRose is wrong; it may well be that one or more of these factors can affect people's salience judgments in particular, or contextualist intuitions more generally. More, and more subtle, tests might reveal that they do. What the contrast between Hansen and Chemla, and Alexander et al. does show is that the *null* results are *not* merely the product of an unfair presentation of the context-shifting cases.

Perhaps more importantly, DeRose should welcome this result in the light of his larger attempt to explain the grip of skeptical arguments in terms of the context-sensitivity of "know." Skepticism is by common accord familiar, and the arguments for it are purported to be intuitive. Like the sorites paradox, it's an easy puzzle to explain and a bedeviling one

[58] DeRose, 2011: 83 - 87.

† I'll just note that while Schaffer and Knobe describe their study as only manipulating salience, as I've argued above, they do introduce stakes inadvertently.

to answer. Taking the common accord as read for an instant, and if we accept DeRose's diagnosis that the paradox depends on manipulating epistemic standards contextually, then it better not be the case that the manipulation in question depends on delicate calibration. If explaining external world skepticism really depended on these kinds of features, it's unlikely anyone would ever have been moved by it in the first place. Most presentations of skeptical arguments don't conform to all of these suggestions, which suggests we should look elsewhere to understand salience effects.

3.4 New Experiments 1a - 1d.

Methods:

One worry about relying on the Alexander et al. results is that their experiments use a single cover story about shopping in a furniture store and it is entirely possible that the results they report are tied to that particular cover story. To demonstrate that salience effects are general, and don't depend on the peculiarities of a given story, it's important to show that a variety of different cases can generate them. The following experiments test this possibility.

To further investigate whether the salience effects are genuine and general, and if they are truly fragile effects that depend on presentation, 211 participants from the United States (Age M=30.1; Gender: 35% Female; Education: 46% Bachelors degree or above) were recruited online from Amazon Mechanical Turk and compensated for their participation at minimum wage. Participants were divided into four conditions using an online randomizer. Each condition used a different cover story: the first was the furniture store vignette from Alexander et al., the second involved identifying animals at the zoo, the third involved identifying plants at an arboretum, and the fourth involved identifying birds on a lake. Each condition presented participants first with a normal story, where no error had been made salient, and followed it with the same story, where an unrealized error possibility was made salient. Following the methodology of Alexander et al., after reading each vignette, participants were asked to rate their agreement to the following statement "John knows the (table/bird/animal/plant) is (red/a Siberian grebe/a jaguar/a gorse bush)" on a six point likert scale where 1 = strongly disagree and 6 = strongly agree. On a subsequent page, with the vignette no longer visible, participants answered a question designed to check their attention: eight respondents were eliminated for failing this question. Participants were then

asked to fill in a standard demographic questionnaire. The new cover stories are modeled on Fred Drestke's Siberian grebe/gadwell duck and zebra/painted mule thought experiments.[59] The cover stories and results are presented in Table 8, and the results are visualized in Figure 3.

	Normal	Salient Error
Furniture Condition	John A. Doe is in a furniture store. He is looking at a bright red table under normal lighting conditions. He believes the table is red.	Normal Case + However, a white table under red lighting would look exactly the same to him, and he has not checked whether the lighting is normal, or whether there might be a red spotlight shining on the table.
Zoo Condition	John and his friend go to the zoo. As they walk around, they pause in front of an exhibit marked "Brazilian Jaguar Enclosure". John and his friend read about jaguars from the sign, and look out and see a jaguar sleeping on the branch of a tree in the enclosure.	Normal Case + However, African leopards look very much like Brazilian jaguars, and the signs in the zoo have recently been replaced by an inexperienced crew of workers. If a zoo official had accidentally switched the signs on their exhibits, John wouldn't have been able to tell the difference between the jaguar he did see and a leopard.
Arboretum Condition	It is a beautiful Saturday early in spring, and John goes to the local arboretum to see the plants flowering and to enjoy the weather. As he walks around, he pauses in front of a yellow flowering bush. The sign next to the bush is labeled "Gorse." In his guidebook he reads that gorse is native to Scotland, and that it is a genus of flowering plants in the family Fabaceae.	Normal Case + However, nearby in the arboretum there is an exhibit of an Asian forsythia which looks remarkably similar to gorse. If the arboretum staff had mixed up the signs, John wouldn't have been able to tell the difference between the gorse bush he did see and a forsythia bush.
Siberian Grebe Condition	John is an outdoor enthusiast. He hears on the radio that a very rare breed of bird, the Tufted Siberian grebe, is roosting in the area, and can be seen swimming on a nearby lake. The radio report describes the Siberian grebes as having black top feathers and white wing-tip feathers. John is an avid photographer, so he drives out to the lake, sees a Siberian grebe, and takes a picture of it.	Normal Case + However, the radio report doesn't mention that gadwell ducks are very common in the area, and that gadwell ducks are almost indistinguishable from Siberian grebes. John is not a birdwatcher, and if the bird had been a gadwell duck, he wouldn't have been able to tell the difference.

Table 8. Experiments 1a - 1d: Participants were asked to rate their agreement to the following sentence "John knows that the (table/animal/plant/bird) is (red/a jaguar/a gorse bush/a Siberian grebe)" on a six point likert scale, where 1 = strongly disagree and 6 = strongly agree.

[59] See F. Dretske (1981: 369) for his Siberian grebe thought experiment, and (1970: 1015-16) for the Zebra/painted mule thought experiment.

Results:

The experiment demonstrates that salience effects are general, and they are not fragile. As Figure 3 clearly shows, participants attribute knowledge at a high level when no error is salient, and at a low level when an unrealized possibility of error is salient.

A Shapiro-Wilk Test was significant at the $p<.000$ for knowledge attribution in the normal case in all four conditions (Furniture store $W=.688$; Arboretum $W=.677$; Zoo $W=.578$; Siberian Grebe $W=.892$). A Shapiro-Wilk Test was significant at the $p<.01$ for knowledge attribution in the skeptical case in all four conditions (Furniture store $W=.929$; Arboretum $W=.930$; Zoo $W=.927$; Siberian Grebe $W=.828$). Consequently, all results were analyzed non-parametrically.

Collapsing across all conditions, and comparing results from the normal case and the error case, we find that participants ($N=203$) attribute knowledge in the normal case ($M=5.14$, $SD=1.14$) at a much higher level than in the error case ($M=3.21$, $SD=1.48$). A Wilcoxon Rank Sum Test showed the difference was statistically significant, $Z=-11.27$, $p<.000$, $r=0.79$ (a "very large" effect).

To ensure that no single cover story was driving the effect, I analyzed each pair separately. Looking first at the Furniture Store condition ($n=56$) knowledge attribution in the normal case ($M = 5.20$, $SD = 1.26$) was higher than in the error case ($M=3.54$, $SD=1.35$). A Wilcoxon Rank Sum Test showed the difference was statistically significant, $Z=-5.61$, $p<.000$, $r=0.75$ (a "very large" effect).

Looking next at the Arboretum condition ($n=56$) knowledge attribution in the normal case ($M=5.25$, $SD=1.06$) was higher than in the error case ($M=3.71$, $SD=1.46$). A Wilcoxon Rank Sum Test showed the difference was statistically significant, $Z=-5.89$, $p<.000$, $r=0.787$ (a "very large" effect).

Turning to the Zoo condition ($n=45$) knowledge attribution in the normal case ($M=5.49$, $SD=0.90$) was higher than in the error case ($M=3.33$, $SD=1.46$). A Wilcoxon Rank Sum Test showed the difference was statistically significant, $Z=-5.37$, $p<.000$, $r=.798$ (a "very large" effect).

Looking finally at the Siberian grebe condition ($n=46$) knowledge attribution in the normal case ($M=4.61$, $SD=1.13$) was higher than in the error case ($M=2.09$, $SD=1.07$). A Wilcoxon Rank Sum Test showed the difference was statistically significant, $Z=-5.76$, $p<.000$, $r=0.850$ (a "very large" effect).

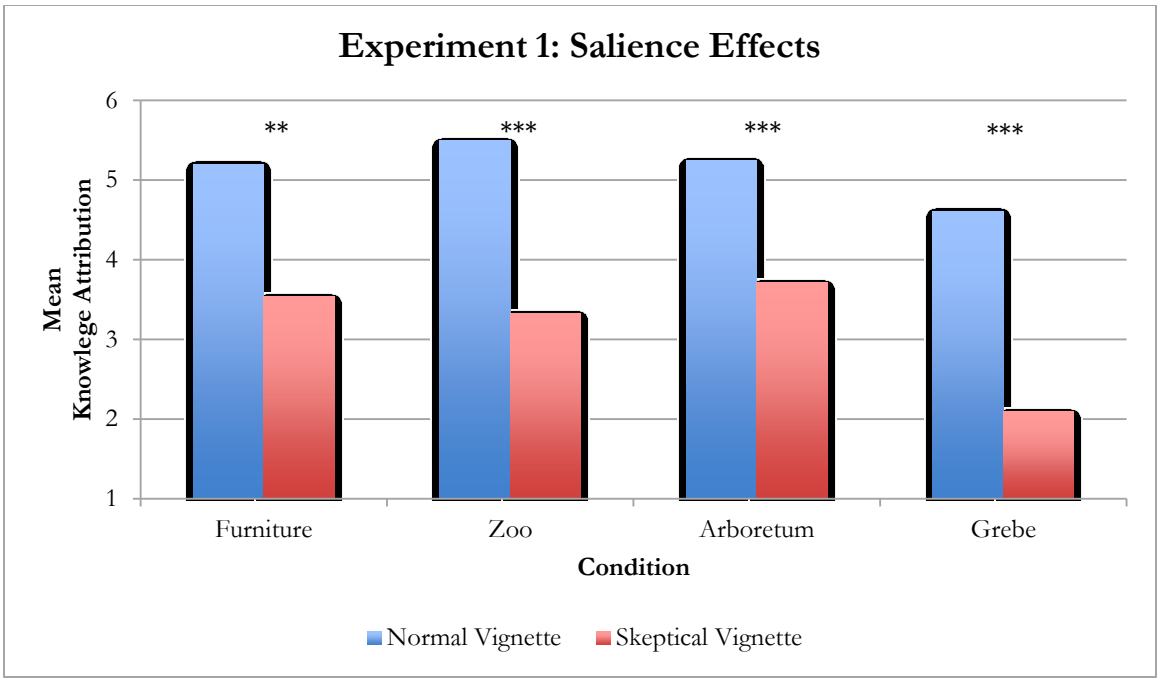


Figure 3. Experiment 1: Salience effects with different cover stories. Significant differences are marked with * where p<.000.**

Discussion:

The results of this study provide powerful support for the contention that salience effects are genuine and general. Looking across four different kinds of cover story, we find that mentioning an unrealized possibility of error drives down knowledge attribution, and that the effect is, statistically speaking, very large. These results also suggest that salience effects are not fragile: they do not, as DeRose hypothesizes, delicately depend on the phrasing of the evaluated assertion, and this suggests that context effects are not dependent on something like Lewis' rule of accommodation. A demographic analysis also suggests that salience effects are not an artifact of individual differences. There were no significant correlations between knowledge attribution in the normal condition for gender (p=.199), age (p=.525), race (p=.390), or education (p=.160). There were no significant correlations between knowledge attribution in the skeptical condition and age (p=.616), race (p=.167), or education (p=.902). There was a weak but highly significant correlation r_s (203)=.214, p<.002 between gender and knowledge attribution in the skeptical condition, with women (n=70) attributing knowledge (m=2.77) at a much lower level than men (n=133) attributed knowledge (m=3.44). This will bear watching in future studies, but it isn't responsible for

salience effects, as both men (Z=-8.75, p<.000, r=.760 a "very large" effect) and women (Z=-7.13, p<.000, r=.857 a "very large" effect) attribute knowledge at significantly lower levels in the error case than in the normal case.

This study aimed at trying to detect whether salience effects are genuine. It did not aim at diagnosing why some cases work, and others don't. Why, for instance, have DeRose's bank cases been so philosophically influential, when they simply don't generate salience effects among the folk? Understanding this question has important implications for the diagnosis of skepticism. If some cases work, and other cases do not, we cannot simply assume that raising an unrealized possibility of will undermine our willingness to attribute knowledge. Once we recognize this, we can see clearly the critical and unargued assumption at the heart of the contextualist's diagnosis of the skeptical problem: that raising a far-fetched possibility of error is just as powerful as raising a quotidian possibility of error. But now that it is clear that some cases work, and others don't, it will be important to identify *why* some cases work, and others don't. I turn to this question in the next section.

4. Salience & Plausibility

If methodology doesn't explain the difference between the null and positive results, then what does? Here we get to heart of the question of salience. Most contextualist discussions are guarded when it comes to proposing specific factors that might influence salience effects. Schaffer and Knobe suggest a number of possible factors, and they're as good a starting point as any.

> [T]he possibility of a bank suddenly changing its hours is a **strange** and **improbable** one, and it might not be possible to make people regard it as relevant merely by mentioning it once in the course of a vignette. Perhaps one can only make people regard a possibility as salient by presenting it in a **concrete** and **vivid** fashion.[60]

The first two features are epistemic, while the second two are more matters of presentation. Allow that the following glosses are close to what Schaffer and Knobe had in mind. A possibility is **strange** if it is surprising or hard to understand given a subject's background knowledge and expectations. A possibility is **improbable** if, relative to a specified domain, it occurs relatively infrequently or not at all, given a subject's background knowledge. A

[60] Schaffer and Knobe, 2012: 694. Emphasis added.

possibility is made **concrete** if it is presented as an actual case rather than as a type of case. And, a possibility is **vivid** if it is presented in a detailed or emotionally engaging way.

Strangeness and probability have clear ties to epistemic evaluation, and they are just as clearly related, as many things that strike us a strange do so because they are uncommon or unfamiliar. And yet, the two concepts can come apart as well. Something can be strange, but not seemingly improbable, at least upon reflection. For instance, some salamanders native to Japan are six feet long.[61] That's strange, and certainly rare, but it's hard to say that it's improbable in a statistical sense. Likewise, plane crashes are statistically unlikely, but they are a familiar kind of event. Even though they come apart, it's easy to see how they can serve a kind of heuristic function in epistemic evaluation. As Lewis himself remarks, once you set your mind to it, unrealized possibilities of error are everywhere. It is practically impossible, and certainly a waste of effort for limited and fallible creatures such as ourselves, to try to rule each and every possible error out. Judgments about strangeness and improbability are front-line filters for epistemic efficiency. If a possibility strikes us as strange or seemingly improbable, relative to background expectations, that is a signal that it doesn't deserve as much effort or attention. But the converse is true as well, the familiar and likely events deserve the most attention - and those familiar possibilities of error are the ones we should try to rule out.[62] From the perspective of an evaluator, failure to rule out likely errors will appear to be an epistemic omission, and could well be at the route of salience effects. They're candidates worth exploring experimentally.

Concreteness and vividness are not straightforwardly epistemic concepts. It's possible to present the same event abstractly or concretely, pallidly or vividly. So, from an epistemic perspective it's less clear why they would, or should, have an affect on salience. However, it is possible to see how they could influence a subject's *attention* to a possibility. Stilted journalism goes unread while vivid and engaged reporting wins prizes. When a possibility is described in detail, or in a way that generates an emotional response, then we do tend to pay attention. And this certainly makes sense, as detailed descriptions that are emotionally engaging are signals of importance. If someone else thinks something is important, then that is a signal we should as well.[63] Concreteness can, consequently, serve as

[61] For a similar point, see: http://wtfevolution.tumblr.com/
[62] Think of Lewis' rule of actuality here.
[63] Think of Lewis' rule of conformity here.

an indirect signal either of practical importance, and thus stakes, or of probability, and thus salience.

Interestingly, Schaffer and Knobe choose to manipulate concreteness and vividness, rather than a more straightforward epistemic property. Their choice makes perfects sense from the perspective of the Lewisian rule of attention. It may well be that the previous studies just failed to draw enough attention to the error possibility that they mention - the survey-makers equivalent of mumbling in conversation. This is good experimental practice. However, in understanding the historical grip of skeptical hypotheses, it would be surprising if their force really depended on matters of presentation. The enduring grip of skepticism suggests it is unlikely to be so fragile.

One reason to focus on plausibility is that there is a tension in the standard contextualist line in how to deal with doubt and the probability of a possibility of error. Contextualists are cagey here. Their diagnosis is premised on the idea that skeptical arguments are intuitive, and therefore it cannot be too taxing for the skeptic to make skeptical hypotheses relevant. Some invariantists like Timothy Williamson and John Hawthorne argue that the implausibility of skeptical hypothesis plays an essential role in understanding the grip of skeptical arguments. They claim that it is because skeptical hypothesis are psychologically arresting that we overestimate the threat to they pose to ordinary knowledge. Both appeal to what's known as the "availability heuristic" in support of this diagnosis. But, as we'll see, both these positions are at odds with commonsense epistemic practice and with a number of recent studies in the psychology of law and politics, which show that introducing implausible error possibilities does not undermine individual's willingness to attribute knowledge. In what follows I'll first look at contextualist and invariantist discussions of plausibility. I'll then review the results of a broad range of psychological studies that show our ordinary epistemic practice is sensitive to the plausibility of an error possibility, and that we tend to ignore implausible error possibilities; I'll call this indifference to implausible error possibilities. Finally, I'll present a series of studies I've conducted which show that plausibility does matter to the efficacy of salience effects, and that implausible error possibilities like those employed in global skeptical arguments *do not*, on their own, generate salience effects. These results are at odds with the invariantist diagnosis, and suggest availability is not at factor in assessing skeptical hypotheses. These results also suggest that the contextualist diagnosis is incomplete. If we typically ignore

implausible error possibilities, then a genuine diagnosis of the skeptical problem must explain how the skeptic does make their global error possibilities threatening. Looking forward, in section 6, I'll present the results of a further set of studies that reveal one mechanism by which the skeptic *can* overcome our ordinary indifference to implausible error possibilities.

4.1 Contextualism and Plausibility

Among contextualists, David Lewis' discussions of discourse constraints inform much of what's said about how the skeptic is able to alter the context of evaluation in their favor. A simple tension faces contextualism: we are not always concerned with the skeptic, nor are we always worried when skeptical hypotheses are aired. So the contextualist must offer an account that is permissive enough to explain how the skeptic can become relevant to a conversation, but one that is restrictive enough to explain why we often rightly ignore conspiracy theories (and Fox News).

In "Scorekeeping in the Language Game" and in "Elusive Knowledge," Lewis attempts to navigate this tension by developing i) a context-sensitive semantics for "know," and ii) a set of pragmatic rules that govern how contexts are determined.[64] In Lewis's early work, he seems to think that it is relatively easy for the skeptic to make skeptical hypotheses relevant. If this is right, than the implausibility of skeptical hypotheses shouldn't be an important factor in understanding salience effects. In "Scorekeeping" he posits that this is in part the result of the rule of accommodation, whereby the set of contextually relevant alternatives shifts in such a way to make a given assertion acceptable in a context.

> The commonsensical epistemologist says: "I know the cat is in the carton - there he is before my eyes - I just can't be wrong about that!" The sceptic replies: "You might be the victim of a deceiving demon". Thereby he brings into consideration possibilities hitherto ignored, else what he says would be

[64] See Ichikawa (2011: 86 - 87) for a similar presentation. Ichikawa tries to implement Lewis' contextualism in terms of a context-sensitive domain restriction for "know." He deflects most of the criticism that's been leveled at contextualism as a criticism of claims that belong to set ii). This kind of defense is a standard move for contextualists, and it preserves their semantic thesis by undermining the epistemic project they're also undertaking. You can say as much as you like that pragmatics are complicated, or that explaining ii. is an "ambitious meta-semantic project" that's destined always to be incomplete as much as you like, but that's not an *explanation* of the grip of skepticism; that's a description of the fact that *something complicated happens* when we're gripped by skeptical arguments. This kind of approach prioritizes the solution they want, getting ordinary knowledge attributions coming out true via contextual case restriction of some kind, over an actual explanation of why such a solution is necessary.

> false. The boundary shifts outward so that what he says is true. Once the boundary is shifted, the commonsensical epistemologist must concede defeat...We get the impression that the sceptic…has the last word. Again this is because the rule of accommodation is not fully reversible. For some reason, I know not what, **the boundary readily shifts outward if what is said requires it,** but does not so readily shift inward if what is said requires that.[65]

In Lewis' later work, he posits a version of contextualism whereby an agent only knows that *p* if their evidence rules out the relevant alternatives to *p*. But what's relevant can vary according context for Lewis. Of the many rules Lewis proposes for governing relevance, and the key to understanding salience effects on his view, is the rule of attention. According to Lewis, if we are attending to a possibility, then that possibility is relevant, and must be eliminated if an agent is to count as know that p. He says,

> Our final rule is the Rule of Attention. But it is more a triviality than a rule. When we say that a possibility is properly ignored, we mean exactly that; we do not mean that it could have been properly ignored. Accordingly, a possibility not ignored at all is ipso facto not properly ignored. What is and what is not being ignored is a feature of the particular conversational context. **No matter how far-fetched a certain possibility may be**, no matter how properly we might have ignored it in some other context, **if in this context we are not in fact ignoring it but attending to it, then for us now it is a relevant alternative**.[66]

According to the rule, once an unrealized error possibility is introduced into a context, and an individual attends to it, then it must be ruled out for the individual to have knowledge. Now, *mentioning* a possibility in a given context may not mean that a subject is *attending* to it; so, for all Lewis has said in explaining the rule of attention, he hasn't committed himself to the overly simplistic idea that merely mentioning a skeptical scenario is enough to automatically alter the conversational context in the manner of a mindless, causal process. One gloss on this is that *ignoring* a defeater must be distinguished from *being aware* of it, so while a participant evaluating a vignette has read the defeater, and thus aware of it, they still might be ignoring it.

[65] Lewis, 1979: 355. Emphasis added.
[66] Lewis, 1996: 559. Emphasis added.

Many contextualists are keen to note this because although skeptical arguments can be gripping, they aren't always, and they don't grip everyone on every occasion.[67] The salience mechanism must be flexible. That said, if what counts as *attending* to a defeater is just being influenced by a defeater, then the rule adds very little to our understanding of how the skeptic threatens ordinary knowledge.

In some of the experiments I've been considering, especially the null results of Buckwalter, May et al., and Hansen and Chemla, the participants still attribute knowledge even though the agents in the stories *had not* ruled out the defeater. This raises a central question: what explains the attribution of knowledge here? Were the participants ignoring the defeaters, even though they were aware of them? Or, were the participants attending to the defeaters, and those defeaters were nevertheless insufficient in some way to undermine their ascriptions of knowledge to the agents in the vignettes?

And this is worth exploring because it is the key feature to which contextualist appeal in explaining skepticism. We cannot take it for granted that if the pattern of shifting intuitions identified by contextualists actually occurs that it will also occur with skeptical defeaters. Evil demons and the like are unfamiliar, strange, and very improbable. The contextualist pattern does exist with ordinary error possibilities, but it doesn't follow that it will reappear with implausible ones - indeed it is worth remembering that contextualists *motivate* their view with ordinary, run-of-the-mill cases, and then *extend* it to skeptical cases. Indeed it is critical to their diagnosis that they don't take the skeptical case as evidence for their position in the first instance. And attending closely to Lewis here, he says it does not matter if an unrealized possibility of error is implausible, if it's not ignored it is relevant. He does not say that plausibility is a factor in whether or not an unrealized possibility will be ignored. But as we saw in Chapter 1 most contextualists do think that the skeptic's argument is intuitive.

---

[67] DeRose, 2009: 138. For reasons I explain in Chapter 5 in more detail, I'm not hopeful for the prospects of DeRose's attempt to explain resistance to "skeptical maneuvers" in terms of conversational moves that tend to "resist" attempts to raise standards or just generally tend to lower them. Quite generally, these are *conversational* conventions, and they lead to confusion and conflict of a kind similar to, but not the same as, the skeptical paradox because the agents involved have different, conflicting conversational intentions. I think the use of conversational mechanisms misplaces what needs to be understood. Skepticism isn't a problem that arises in earnest debate with another individual, but in one's self. The means and mechanisms of conversational moves are just misplaced in what is essentially a personal matter. See Ichikawa (2011) for similar demurrals regarding the rule of attention. For my part, what I find implausible about this style of response is that it proposes that a single individual has two conversational intentions, and they can recognize neither.

But here we get to the crux of a difficult question. Contextualists follow Lewis in allowing that the skeptic isn't always successful, but they must allow that the skeptic is sometimes, if not generally, successful, and that the skeptic takes advantage of our ordinary epistemic practices, otherwise the materials from which they assemble their diagnosis would no longer fit the subject of that diagnosis. There is a definite tension here.[68] On the one hand, they must explain why we ignore the skeptic some times, just as we ignore the fanciful and the far-fetched. But, on the other, contextualists must allow that our ordinary epistemic practices are relatively skeptic-friendly, that it is generally fairly easy for the skeptic to manipulate the context in their favor. This fits with much of what is said about the intuitiveness of skeptical arguments, though there is certainly a range of opinions. Some tend to treat plausibility as an inessential feature of epistemic evaluation. For instance, in discussing his famous zebra/painted mule cases, Fred Dretske is clear that plausibility doesn't matter,

> Granted, the hypothesis (if we may call it that) is not very plausible, given what we know about people and zoos. But the question here **is not whether this alternative is plausible, not whether it is more or less plausible than that there are real zebras in the pen,** but whether you know that this alternative hypothesis is false. I don't think you do. In this I agree with the skeptic.[69]

David Lewis, once he turns his attention away from the discussion of the rule of attention, also seems comfortable with extending the contextualist machinery to implausible sources of error, saying,

> Let your paranoid fantasies rip--CIA plots, hallucinogens in the tap water, conspiracies to deceive, old Nick himself--and soon you find that uneliminated possibilities of error are everywhere. **Those possibilities of error are farfetched**, of course, but possibilities all the same. **They bite into even our most everyday knowledge**.[70]

---

[68] Intriguingly, Lewis (pg. 561) acknowledges that some people might be steadfast in their resistance to the skeptic, and both attend to the skeptical hypothesis, but resist the idea that it is relevant. He says, "If you say this, we have reached a standoff. I started with a puzzle: how can it be, when his conclusion is so silly, that the sceptic's argument is so irresistible? … If you continue to find it eminently resistible in all contexts, you have no need of any such explanation. We just disagree about the explanandum of the phenomenon." He may be right that there is not a single explanadum, but two: one for ordinary speakers, and another for those who have something in common with philosophers. For my part, I believe that a full diagnosis will include an account that appeals to features of i) skeptical hypotheses (plausibility), ii) context (previously introduced skeptical hypotheses) and iii) individual differences (need for closure and cognitive reflection tendencies).

[69] F. Dretske, 1970: 1015-16 (emphasis added).

[70] D. Lewis, 1996: 549.

And when this stance is combined with his infallibilist account of knowledge, it seems to yield a particularly permissive form of contextualism. Lewis argues in favor of infallibilism linguistically, saying, "To speak of fallible knowledge, of knowledge despite uneliminated possibilities of error, just sounds contradictory."[71] The power of infallibilism is in explaining how the patently improbable, even delusional possibilities of evil scientists and deceiving demons can underwrite Cartesian skepticism. And yet, at the same time, contextualists often recognize that this cannot be quite right. Sometimes the blatantly implausible is just ignored. If it weren't, the only task of a defense lawyer would be drawing up paranoid fantasies.[72] But how and why is less clear. Lewis suggests that "make-believe" ignoring can ripen into real ignoring.[73] And while he's certainly right that there is a voluntaristic aspect to attention, it can't be purely voluntaristic: while we may choose to study the skeptical problem, we do not choose to be troubled by it.

Keith DeRose treads carefully in addressing this tension. Though he acknowledges that radical skepticism rests on implausible hypotheses, he nevertheless emphasizes how they threaten ordinary knowledge,

> [H]owever improbable or even bizarre it may seem to suppose that I am a BIV, it also seems that I don't know that I'm not one. How could I know such a thing?[74]

Elsewhere he acknowledges what he calls the "Aw, come on" response when skeptical arguments like AI are introduced, saying that many students find those arguments to be unconvincing.

> [Some students think AI is] farfetched, ridiculously weak, and quite unthreatening; such a reaction is often accompanied by an exclamation somewhat along the lines of, "Aw, come on!" Those inclined to react in this latter way probably grew increasingly impatient of my repeated description, in section 1, above, of the Argument by Skeptical Hypothesis as "powerful," thinking instead that the argument hasn't a chance in the world of establishing its absurd conclusion.

---

[71] Lewis, 1996: 549.
[72] *Ibid*: 560.
[73] *Ibid*.
[74] DeRose, 1995: 2-3. (emphasis added)

> Well, the skeptical argument really is powerful -- at least fairly powerful -- and is certainly not absurdly weak. The argument is clearly valid -- its premises imply its conclusion -- and each of its premises, considered on its own, enjoys a good deal of intuitive support. **For however improbable, farfetched, or even bizarre it seems to suppose that you are a brain-in-a-vat,** it also seems that you don't know that you're not one**.** How could you possibly know such a thing? And it also seems that if, for all you know, you are a brain-in-a-vat, then you don't know that you have hands. How could you know you have hands if, for all you know, you're bodiless, and therefore handless?[75]

Indeed, it's not clear here if DeRose is really committing himself to the idea that plausibility doesn't matter, or attempting to overcome it by putting the burden of ruling out the skeptical hypotheses on those with the 'aw shucks' response. Gail Stine, an early architect of contextualism, also addresses the relation of the plausibility of skeptical hypotheses and our ordinary epistemic practice, though indirectly, when she says,

> It is an essential characteristic of our concept of knowledge that tighter criteria are appropriate in different contexts. It is one thing in a street encounter, another in a classroom, another in a court of law—and who is to say it cannot be another in a philosophical discussion? . . . We can point out that some philosophers are very perverse in their standards (by some extreme standard, there is some reason to think there is an evil genius, after all)—but **we cannot legitimately go so far as to say that their perversity has stretched the concept of knowledge out of all recognition—in fact they have played on an essential feature of the concept.**[76]

In a recent paper evaluating the extant linguistic analogies that contemporary contextualist have explored as possible models for a context-sensitive semantics for "know," Jonathan Schaffer and Zoltán Gendler Szabó adopt as a desideratum of any such account the it explain how it is relatively easy for the skeptic to expand the set of relevant alternatives in a context, and how it is difficult to contract them. Schaffer and Szabó are relatively pessimistic about the prospects, but they nevertheless note that,

> [I]t is part of the standard contextualist treatment of skeptical arguments that there are overall discourse constraints that **make it easy to expand the relevant alternatives**, but difficult to contract them once they are introduced into the discourse.[77]

---

[75] DeRose, 1999. (bold emphasis added)
[76] G. Stine, 1976: 254. (emphasis added)
[77] J. Schaffer & Z. G. Szabó, 2014: 501.

It would seem, then, that contextualists of various types discount plausibility in epistemic assessment. They are somewhat elusive when it comes to explaining if plausibility matters, but in general they *do not* see an outstanding need to offer an explanation of why plausibility shouldn't matter, or indeed to extend their diagnosis to an explanation of how the skeptic overcomes a commonsense resistance to entertaining the implausible. But this deserves exploration, because any adequate diagnosis of the skeptical problem will shed light on how patently implausible skeptical hypotheses sometimes get a grip on us, but delusional conspiracy theories rarely do. That many - perhaps most - people have the "Aw, come on" response to radical skeptical hypotheses is a *feature*, not a *fault* in our response. It is a feature that must be part of an empirically respectable diagnosis of the skeptical problem.

4.2 Invariantism and Plausibility

Some epistemologist do think that plausibility matters, and far from thinking that it poses a problem for diagnoses of the skeptical argument, they believe it is essential to understanding the power of the skeptic's threat. Timothy Williamson and John Hawthorne are anti-skeptical invariantists who accept the pattern of intuition shifts that contextualists point to, but argue that these shifts are not genuine evidence for contextualism. Instead, they argue that skeptical hypotheses are psychologically gripping because they are fanciful, and this leads us to overestimate the epistemic threat they pose. Williamson contends,

> One effect of fictional violence on television is to make viewers overestimate their chances of being the victims of violent crime: they suffer an illusion of danger. Might not an illusion of epistemic danger result from exposure to lurid stories about brains in vats, evil demons, painted mules, or gamblers who bet the farm?[78]

John Hawthorne concurs, offering a very similar diagnosis of the origin of salience effects,

> When certain non-knowledge-destroying counter possibilities are made salient, we overestimate their real danger; as a result, we may find ourselves inclined to deny knowledge to others in cases where there is in fact no real danger of error.[79]

[78] T. Williamson, 2005: 226.
[79] J. Hawthorne, 2004: 164.

Both account for this effect in terms of a cognitive bias Daniel Kahneman and Amos Tversky have dubbed the availability heuristic. According to Kahneman and Tversky we tend to use the ease with which a possibility is imagined as a proxy for determining the likelihood of the event.[80] The classic example of the availability heuristic at work comes from assessments of the danger posed by different forms of travel: people tend to overestimate the danger of flying and underestimate the danger of driving because plane crashes often occasion breathless media attention and graphic visuals. Or, to take another example, people tend to overestimate the danger posed by sharks and snakes, and underestimate the danger posed by unwashed food and the flu.

On this kind of view, it is essential that the skeptic's hypothesis is somehow out of the ordinary, like the thought of a super-psychologist or a malevolent demon, and because it is psychologically arresting, it becomes epistemically arresting too.[81] Appealing to a cognitive bias, and perhaps the most well known one at that, does normative work for Williamson and Hawthorne as well. Biases are rational failures: they are particular circumstances where we are prone to be less than rational, and in a way that we would *acknowledge* if the facts were pointed out to us. Because, according to their account, the force of the skeptical hypothesis depends on a more general bias in our reasoning, the intuitiveness of the skeptical challenge relies on a principle that we do not in general endorse. This offers them a way to resist the idea that the skeptical challenge really does appeal to our ordinary concept of knowledge, and by extension a way to resist the threat that skepticism poses.

As Jennifer Nagel rightly observes, while assessing Williamson and Hawthorne's proposal, the existing research into the availability heuristic doesn't issue a clear verdict when it comes to how ordinary individuals will respond to skeptical scenarios. Kahneman and Tversky originally frame their work in terms of the "ease with which relevant instances come to mind,"[82] and in so far as the skeptic encourages us to consider a hypothesis, it doesn't follow that it is easily imagined in these terms. Indeed, the evidence is mixed. Sometimes encouraging individuals to imagine *vivid* and *familiar* scenarios like plane crashes can result in overestimated likelihood. But some hypotheses are hard to imagine because they are

[80] A. Tversky & D. Kahneman, 1973.
[81] See J. Nagel, 2010, for a critique of the invariantist appeal to the availability heuristic.
[82] D. Kahneman & A. Tversky, 1973: 207.

unfamiliar or described in a way that is vague. In these cases encouraging an individual to imagine them can cause them to discount the likelihood of the possibility.[83]

It is simply unclear whether the extant work really supports a single verdict when it comes to global skeptical hypotheses, or whether skeptical hypotheses would all function in the same way to most people. Might demon scenarios feel familiar to the religious, but unfamiliar to the agnostic? Might envatted brains feel unfamiliar to fans of historical fiction, but familiar to fans of science fiction? I think these are open questions, especially given the difficulty of operationalizing "ease" of imagining.[84]

4.3 Plausibility and the Psychology of Law and Politics

Recent work in the psychology of law and political commitment raises some problems for this invariantist proposal and for forms of contextualism that are permissive towards the implausible. Psychologists of law and lawyers focused on jury selection have long been interested in what factors affect the conception of reasonable doubt. A number of studies have shown that introducing alternative hypothesis to explain events can be a powerful tool for inducing individuals to revise their initials beliefs, or to revise their confidence in their initial beliefs. But across these studies, whether they are of doubt in juries, or of the retraction of misinformation, a critical factor is that the alternative hypothesis is plausible. But this is just commonsense: we are limited epistemic agents, we cannot consider every hypothesis in depth, so filtering those possibilities we attend to based on subjective plausibility given background knowledge is what you'd expect of an efficient system. More interestingly, and contrary to the invariantist and contextualist proposal considered here, some of the studies I'll canvas in this section suggest that introducing implausible alternatives to a default or focal hypothesis can actually cause individuals to increase their confidence in the focal hypothesis.

One suggestive source of evidence about how individuals evaluate rival hypotheses comes from work in the psychology of politics. It is not unsurprising to find, for example, that many people still believe that Saddam Hussein controlled a large research campaign into

[83] See Nagel, 2010: 292 for a review of the relevant experimental work, especially Steven Sherman's work (1985: 120) on estimating the likelihood of contracting diseases when described in a concrete or abstract way. Abstractly described scenarios were rated as less likely than concretely described scenarios.

[84] Other non-skeptical epistemologists believe we should rule out skepticism based on plausibility, including Michael Ayers, Lawrence Bonjour. See M. Devitt, 1991: 75.

weapons of mass destruction, and that WMDs were found in Iraq after the second gulf war. This is despite the fact that a CIA authored report released in 2004, which was widely reported on in the mainstream media, conclusively stated that no evidence of WMDs or WMD programs had been found. Why are some people so prone to misinformation and so unlikely to revise their beliefs? In a wide ranging survey of debiasing and misinformation, Stephan Lewandowsky, Ullrich Ecker, Collen Seifert, Norbert Schwarz, and John Cook found that the most successful attempts at debiasing involved the presentation of a *plausible* alternative explanation of the facts of a case that explained how the initial misinformation was first accepted. As they say, it is critical that the alternative explain, "why the misinformation was thought to be correct in the first place."[85]

How does this relate to salience effects? It offers a general indication of how individuals typically evaluate rival hypotheses. In the typical salience case, there is a default belief held by an agent, which typically is well supported enough to intuitively count as knowledge. When unrealized possibilities of error are introduced, these function as rival hypotheses, that would lead to a false belief, but are otherwise compatible with all the common-ground evidence. So, our default hypothesis is that the animal in the enclosure labeled 'leopard' is a leopard, but this is undermined when the possibility of a misplaced sign is introduced. Likewise, those in the grip of misinformation have a default hypothesis, i.e. that Saddam Hussein had WMDs, which is then challenged by another, i.e. that he didn't. Lewandowsky et al. report that the most effective counter hypotheses are those that offer a plausible explanation of why the original report was accepted, e.g. that a chemical factory (rather than a farm or an amusement park) was mistaken for a WMD factory (rather than simply being made up).[86] So, poor or implausible rival hypothesis are far less effective at modifying individuals' assessments of a default. The cases are different from a standard salience case in some ways. In the standard salience case it's stipulated that the agent has a true belief, but that they haven't ruled out some possible ways in which their belief could have turned out false. Lewandowsky et al., by contrast, are actually focused on attempts to revise agents' false beliefs. From a more general perspective, however, the work they canvas suggests that plausibility matters and that implausible rival hypothesis tend to be discounted.

[85] S. Lewandowsky, U. Ecker, C. Seifert, N. Schwarz, & J. Cook, *forthcoming*: 36.
[86] *ibid.*

Work in the psychology of law also sheds light on how individuals assess rival hypotheses of differing plausibility. One intriguing study of the role of rival hypotheses on jury deliberation lends support to the idea that salience effects are a general feature of epistemic assessment. Elizabeth Tenney, Hayley Cleary, and Barbara Spellman investigated the efficacy different defense strategies. One strategy is to contest the prosecution's evidence, and argue that it doesn't establish beyond a reasonable doubt that a defendant is guilty. A second strategy is to induce doubt in the jury by arguing that the prosecution's evidence is consistent with a different individual having committed the crime at issue. They found that when the defense offers a plausible alternative suspect, also known as a "plan B suspect," it decreases the likelihood that a jury will find a defendant guilty. The strategy, familiar from shows like *Matlock*, is to defend one suspect by simply offering an account of how another individual had the motive and means to commit the crime in question. In their studies they asked undergraduates acting as mock juries to read fictional trial transcripts where an agent is accused of murder. In some conditions the transcripts only included evidence against the defendant; in others it included an attempt by defense attorneys to show that the prosecution's case was incomplete; and in other conditions the defense included 1, 2, or 3 plan B suspects. Tenney et al. were careful to construct the transcripts so that both the roles of the defendant and of the plan B suspect were consistent with all the evidence presented.[87]

Tenney et al. found that the inclusion of a plan B suspect significantly decreased the likelihood that an individual would give a guilty verdict. They also found that the inclusion of a plan B suspect increased the confidence of those that gave a not-guilty verdict, and decreased the confidence of those that did give a guilty verdict. Plan B suspects are like unrealized possibilities of error: they are possibilities that are consistent with all the evidence. While there may be a default or focal hypothesis, in this case that the defendant is guilty, introducing a hypothesis consistent with evidence decreases confidence in that hypothesis. Tenney et al. found that adding additional plan B suspects did not further decrease the likelihood of guilty verdicts. Participants still tended to think the defendant was the most likely, but so long as the plan B suspect was plausible, it undermined the jurors' confidence enough to reduce guilty verdicts significantly. They conclude, "If employing the [plan B]

[87] E. Tenney, H. Cleary, & B. Spellman: 2009.

defense, one should be careful to construct *plausible* alternative stories…in our experiment, none of the [plan B] suspects were ever thought to be guilty. Exactly what constitutes "enough" plausibility remains to be seen, and is likely to depend on the case-specific circumstances."[88]

Tenney et al.'s results are quite striking, and appear to be a judicial extension of salience effects. While they're careful to say on a number of occasions that it is important that the plan B suspects be plausible as suspects, they do not manipulate the plausibility of the rival hypothesis directly. However, Craig McKenzie, Susanna Lee, and Karen Chen have conducted a number of studies that do just this. To understand how juries respond to evidence and update their confidence in an explanation of events, McKenzie et al. conducted a series of experiments where they asked participants to assess their confidence that a defendant was guilty, while varying the strength of the evidence offered in support of the defendant.[89] They found that when a weak defense, where very little evidence is adduced in support of the defendant's innocence, was presented after a strong prosecution, participants found that the prosecution's case was significantly more convincing than when the prosecution's case was considered alone. In essence they asked participants to rate their confidence after considering two rival hypotheses. Interestingly, they found the same pattern of results both in legal scenarios and in interpersonal disputes, as when a politician is accused of some nefarious activity (a default hypothesis), and then offers an alternative account of their behavior (the rival hypothesis). When the rival hypothesis was implausible, participants were more confident that the default hypothesis was true than when the default was considered on its own. Far from exerting a small downward pressure on our confidence in a default hypothesis, implausible rival hypothesis can make the default seem even better. McKenzie and his colleagues sum up their results in the following way: "Four experiments showed that cases independently rated as weakly supporting one side often increased confidence in the opposing side. The implicit reasoning seems to be along the lines of 'If that's the best they can do, then I believe the other side (even) more'."[90]

Again, there are important disanalogies between this research paradigm and the skeptical arguments I've been considering. First, McKenzie et al. focused on confidence, not

[88] E. Tenney, H. Cleary, & B. Spellman, 2009b: 37.
[89] C. McKenzie et al., 2002.
[90] C. McKenzie et al., 2002: 15.

whether the participants acting as jurors really *knew* that one of the alternative accounts of events was correct. As knowledge and belief are different propositional attitudes, it is certainly open that our ordinary practices of assessment are quite different. Indeed, courts of law do engage with different epistemic standards. Certainly if you know that a defendant is not guilty, you should vote to acquit, but the standard in play in criminal law is of being "beyond a reasonable doubt," and if a defendant's version of events does not induce a reasonable doubt of their guilt, a jury can rightly vote to convict even though their strong belief falls short of knowledge. When the standard is unpacked for juries, various definitions are sometimes given, but among the most common are "having an abiding conviction," and having "moral certitude,"[91] and these strict standards should generally be friendly to skepticism.

Nevertheless, there are two reasons for thinking these results are relevant. First, McKenzie et al. used cases where the defendant had little to no evidence to support their version of events. Just like the radical skeptic, they presented merely possible alternative explanations. We have no positive reason to think we are globally deceived, we just can't rule it out. Second, together with the studies I've just examined, these results are indicative of a more general epistemic strategy when two competing hypotheses are in play: we compare them, and implausible rival hypotheses tend to do little to undermine confidence in the default hypothesis. It's worth investigating if this strategy is more general.

In another investigation of likelihood and confidence judgments, Paul Windschitl and John Chambers found evidence for a contrastive approach to evaluating evidence, whereby including weakly supported or unlikely outcomes can increase the perceived likelihood of a better supported outcome. They call this the "dud-alternative" effect. As they recount, a normative model of subjectivity probability assessments would predict that when comparing a focal outcomes (or default hypothesis) to alternatives, how the evidence for alternative outcomes is distributed should not have an effect on people's perceptions of the certainty of a focal outcome. For example, a person's intuitive certainty of winning a raffle should be indifferent between the following distributions, one where they held 12 tickets and others held 10 and 8, and a raffle where they held 12 tickets, and the others held 1, 2, 1, 3, 2, 1, 4, 1, 2, and 1. What Windschitl found, however, is that people feel much

[91] I. A. Horowitz, 1997.

more certain that they'll win the latter lottery than the former even though the combined odds are the same. As they summarize, "adding very weak alternatives—*duds*—to a likelihood question can increase the judged likelihood of the focal outcome."[92]

In another experiment they presented participants with a fictional survey of favorite foods, and they asked the participants to rate how likely they thought that a focal outcome would be the most frequently selected option in a survey of school-aged children. In one condition without any duds, the list included pizza (the focal option) and hamburger; the second condition included duds like eggplant parmesan and grilled fish along with hamburger and pizza. The results were striking: participants predicted that the focal option to be much more likely when duds were present than when they were absent. Thus, the subjectively most likely hypothesis, that kids prefer pizza, seems even more likely when it is contrasted with a subjectively unlikely hypothesis, that grilled fish is tops among tots.

Windschitl and Chambers replicate the result across a number of modalities, and also find evidence of the effect in studies of police line-up construction.[93] When an innocent suspect who fits the description of the culprit is presented in a lineup with a number of distractors who do not, i.e. dud alternatives, then false positives increase dramatically over lineups constructed entirely of individuals who do fit the description.[94] Although the same caveats apply here as before, extending these results to understanding the cognitive significance of skeptical hypotheses, the Windschitl and Chamber's dud-alternative effect is at odds with both the contextualist and invariantist accounts considered above. If we take the familiar and well-supported real-world hypothesis as the focal hypothesis, and we take subjectively implausible skeptical hypothesis as a dud, then we should expect that ordinary practice discounts them entirely. Indeed, the dud-alternative effect suggests that introducing a skeptical hypothesis might even lead to an increased confidence in the real world hypothesis.

Summing up, there is evidence from a broad array of fields and experimental paradigms that when comparing rival hypothesis, the plausibility of a hypothesis is a central component of distributing our confidence in the truth of those hypotheses. Furthermore,

[92] Windschitl and Chambers, 2004: 199.
[93] P. Windschitl and R. Chambers, 2004: 212.
[94] Wells, Rydell, & Seelaua: 1993.

this strategy makes sense for limited epistemic agents such as ourselves: because we have finite capacities, we should not spend time ruling out implausible alternatives.

A perfectly reasonable objection is that *knowledge ascription* is simply different from the other non-knowledge assessments of confidence and likelihood I've canvassed in this section. Knowledge is a high-status epistemic state, and while I might remain extremely confident that the best explanation of my perceptions is the real-world hypothesis, I might nevertheless refrain from ascribing knowledge when subjectively implausible skeptical hypotheses are in play because I can no longer be absolutely certain. Recall Lewis,

> If you are a contented fallibilist, I implore you to be honest, be naive, hear it afresh. "He knows, yet he has not eliminated all possibilities of error." Even if you've numbed your ears, doesn't this overt, explicit fallibilism still sound wrong?[95]

So, while the forgoing empirical work does not settle the issue, it raises a question, does plausibility matter to salience effects? Does knowledge ascription, like confidence assessments, discount implausible alternatives? Or, as Lewis suggests, is knowledge ascription sensitive to implausible alternatives because it has such high standards? The best way to answer that question is to address it directly by experimentation.

5. Testing Salience and Plausibility

5.1 Experiment 2a

Methods:

To investigate whether the plausibility of an error possibility affects knowledge attribution 202 participants from the United States (Age M=31.6; Gender: 43% Female; Education: 44% Bachelors degree or above) were recruited online from Amazon Mechanical Turk and compensated for their participation at minimum wage. Participants were divided into four conditions using an online randomizer, and read one of four cases: a normal case, a plausible error case, an implausible error case, and a very implausible error case.[96] The normal case serves as a control, and no unrealized possibility of error is mentioned; the

[95] *Ibid*: 550. A similar argument is offered by Peter Unger in "A Defense of Skepticism."

[96] The planned size of the study was 200, with 50 participants per condition. Because some exclusions were expected, 210 participants were recruited online, and 8 eliminated prior to analysis for failing a comprehension question. Because the online randomizer does not sort participants into precisely equal groups, the sub-populations of the conditions range from 40 - 56.

plausible error case mentions the possibility of a plausible unrealized possibility of error; the implausible error case mentions the possibility of error that is subjectively less likely; and the very implausible error case mentions the possibility of an error that while logically possible, is subjectively very unlikely. Following the methodology of Alexander et al., after reading the story participants were asked to rate their agreement to the following statement "John knows the bird is a Siberian grebe" on a six point likert scale where 1 = strongly disagree and 6 = strongly agree. On a subsequent page, with the vignette no longer visible, participants answered a question designed to check their attention: eight respondents were eliminated for failing this question. On a further page, with the vignette no longer visible, participants were then asked to rate how plausible they thought the scenario was on a six point likert scale where 1 = very implausible and 6 = very plausible. Participants were then asked to fill in a standard demographic questionnaire. The vignettes are presented, along with mean knowledge attribution rates and plausibility assessments in Table 9 below.

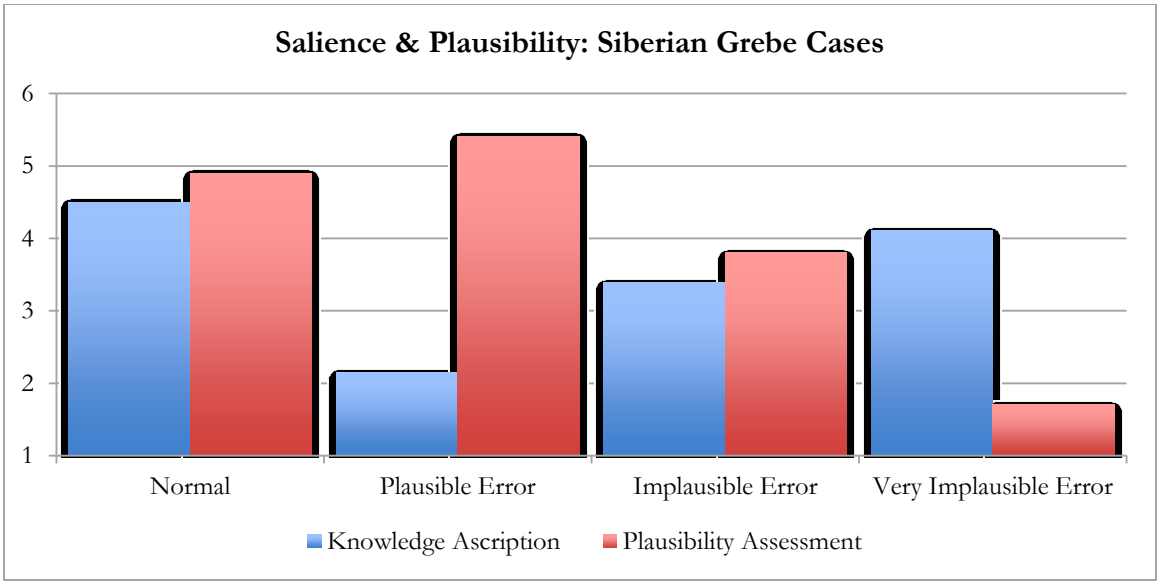


Figure 4: Experiment 2a: Siberian grebe cases. Focusing on the three error cases to the right highlights the relationship between plausibility and knowledge attribution: as plausibility goes down, knowledge attribution goes back up. There is no statistically significant difference between knowledge attribution in the normal case and the very implausible error case.

	Condition	Mean Knowledge Ascription	Mean Plausibility Rating
Normal Case	John is an outdoor enthusiast. He hears on the radio that a very rare breed of bird, the Tufted Siberian grebe, is roosting in the area, and can be seen swimming on a nearby lake. The radio report describes the Siberian grebes as having black top feathers and white wing-tip feathers. John is an avid photographer, so he drives out to the lake, sees a Siberian grebe, and takes a picture of it.	4.50	4.90
Plausible Error Case	Normal Case + However, the radio report doesn't mention that **gadwell ducks are very common in the area**, and that gadwell ducks are almost indistinguishable from Siberian grebes. John is not a birdwatcher, and if the bird had been **a gadwell duck**, he wouldn't have been able to tell the difference.	2.15	5.41
Implausible Error Case	Normal Case + However, the radio report doesn't mention that **it is possible to make mechanical replicas** that are almost indistinguishable from Siberian Grebes. John is not a birdwatcher, and if the bird had been **a mechanical replica**, he wouldn't have been able to tell the difference.	3.39	3.80
Very Implausible Error Case	Normal Case + However, it's possible that **Aliens from another planet could make mechanical replicas** that are almost indistinguishable from Siberian Grebes. John is not a birdwatcher, and if it had been **a mechanical replica made by Aliens from another planet**, he wouldn't have been able to tell the difference.	4.11	1.71

TABLE 9: Salience and Plausibility Experiment 2a. Knowledge ascription means indicate participants response to the following prompt: "Please indicate the extent to which you agree or disagree with the following statement: 'John knows the bird is a Siberian Grebe'." Responses were measured on a six point likert scale where 1 = strongly disagree and 6 = strongly agree.

Results: Experiment 2a

Figure 4 presents a clear picture of the general relationship between the plausibility of an error possibility and knowledge attribution: when a defeater is plausible it is efficacious in undermining ordinary individuals intuitions that an agent in a vignette possesses knowledge. As an error possibility becomes less plausible, it is less efficacious at undermining intuitive knowledge attributions. Because a Shapiro-Wilk Test was significant

for knowledge attribution in all four conditions at the $p<.000$ level (normal, $W=.806$, $p<.000$; plausible error, $W=.788$, $p<.000$; implausible error, $W=.894$, $p=.000$; very implausible error, $W=.875$, $p<.000$) the results were analyzed using non-parametric statistics.

An omnibus Kruskal-Wallis showed a significant effect for condition $H(3)=48.23$, $p<.000$, $\eta^2=.240$. Post-hoc tests showed significant differences between the normal case ($n=40$; $M=4.5$; $SD=1.11$) and the plausible error case ($n=52$; $M=2.15$; $SD=1.42$), $H(1)=74{,}49$, $p<.000$, $\eta^2=.429$ (a "large" effect). This result is consistent with past experimental demonstrations of salience effects. Moreover, it shows that when an unrealized error possibility is considered highly plausible, as in the plausible error case, it can push intuitive knowledge attributions well below the midline of the likert scale. A Kolmogorov-Smirnov test, $D=.225$, $p<.000$, showed a statistically significant difference between actual responses and the midpoint of the scale.

There was also a statistically significant different between the normal case and the implausible error cases ($n=54$; $M=3.39$; $SD=1.74$), $H(1)=-40.32$, $p=.002$, $\eta^2=.097$ (a "small" effect). Notice, however that the effect size in this case is much smaller. Participants rated the defeater in the implausible error case as much less plausible ($M=3.80$; $SD=1.56$), and consequently it was much less effective at generating salience effects.

Most importantly, there was no significant difference between knowledge attribution in the normal case and knowledge attribution in the very implausible error case ($n=56$; $M=4.11$; $SD=1.63$), $H(1)=11.23$, $p=1$, $\eta^2=.006$. Participants rated the plausibility of the unrealized possibility in the very implausible error case as very low ($M=1.71$; $SD=1.37$), and consequently it did not produce a salience effect. Quite strikingly, mentioning a subjectively very implausible unrealized possibility of error is like mentioning no defeater at all! As one would expect then, there was a statistically significant difference between knowledge attributions in the plausible error case and the very implausible error case, $H(1)=-62.25$, $p<.000$, $\eta^2=.285$.[97]

To directly test for a relationship between plausibility assessments and knowledge attributions, I performed a regression of plausibility judgments on knowledge ascription. The regression model use plausibility assessments as a predictor for knowledge attribution in the three cases that mentioned an unrealized possibility of error. There was a significant

[97] Give all pairwise comparisons:

inverse relationship between plausibility rating and knowledge attribution at the p <.000 level; see Table 10.[98]

	B	SE *B*	Beta
Constant	4.73	0.25	
Plausibility rating	-0.415	0.062	-.465*

*r^2=.216, p<.000

Table 10: Regression model for Experiment 2a: Siberian grebe cases.

Discussion:

This experiment has two key results. The first is that there is no difference between people's intuitive knowledge attributions in a case where no defeater is mentioned and a case where a subjectively very implausible defeater is mentioned. Subjectively implausible defeaters, like those that are used to motivate global skepticism do not generate salience effects. The second important result is that plausibility assessment is a significant predictor of knowledge attribution when a defeater is mentioned: when the error possibilities are plausible, salience effects are large; as error possibilities become more implausible, salience effects become smaller and even disappear. These results offer a possible, if surely partial, explanation of the experimental disagreement over the role of salience effects in ordinary intuitions about knowledge: in the null results of Buckwalter and May et al. the error possibility was presented in such a way that participants didn't see it as a plausible epistemic worry, whereas in the results of Alexander et al., it was presented in a concrete enough way to arouse participants epistemic concern. Indeed, these results are consistent with Schaffer and Knobe's suggestion that plausibility might affect salience.[99] Although all of the possibilities they mention (vividness, strangeness, and concreteness) deserve to be explored, these results suggest that the last, plausibility, can have a very large influence. One possible worry about salience effect studies such as these is the possibility that participants deny that agents in the vignettes have knowledge because they misunderstand the case. One could worry that rather than seeing the unrealized error possibility as *just* a possibility, because of

[98] Demographic analysis: there was a significant positive correlation between knowledge attribution and education (r^2_s=.178, p=.012) and an inverse correlation between knowledge attribution and race (r^2_s=-.203, p=.004). These correlations were not replicated in other studies reported later, so they aren't serious concerns as influences in the present study. There were no significant correlations between knowledge attribution and gender (p=.641), age (p=.622), or previous philosophy education (p=.484). There were no significant correlations between plausibility attribution and gender (p=.835), age (p=.548), race (p=.299), or previous philosophy education (p=.824).

[99] Schaffer and Knobe, 2012: 694

shallow processing they think that it is an *actual* error, and that in the vignette the agent's belief is flat false, and thus doesn't count as knowledge. The attention check in the survey was designed to detect this, and asked participants "In the brief story you just read, what did John see? A. a Siberian grebe; B. a gadwell duck (or, mechanical replica of a Siberian grebe, or an illusion of a Siberian grebe); C. a Canadian goose." Answer B was always varied so that it matched the potential defeater from the case. Analysis was restricted to only those participants who answered the attention check correctly, so it is safe to conclude that salience effects are not the product of misunderstanding or shallow processing. It is worth mentioning that there is no evidence of the dud alternative effect here: mentioning a very implausible error case does not drive knowledge attribution even higher. Mean results in the normal case and the very implausible case are not at ceiling, so it is not that the scale won't accommodate an increase. As we'll see, responses to very implausible error cases are categorical: most people ignore them, and because they are at ceiling, there is no room to record a dud alternative effect; a few people do not ignore them, and they do experience salience effects. This small group tends to explain the (non-significant trend) toward lower knowledge attribution in the very implausible case.

However, as in any case, a natural worry is that the inverse relationship between plausibility and knowledge attribution in salience cases is not general, and depends in some way on the specifics of these cases. To explore this possibility, I prepared a follow up study.

5.2 Experiments 2b, 2c, and 2d.

Methods:

To further investigate whether the plausibility of an error possibility affects knowledge attribution 463 participants from the United States (age M=30; gender: 40% female; education: 40% bachelors degree or above) were recruited online from Amazon Mechanical Turk and compensated for their participation at minimum wage. I prepared three variations of the stories presented in Experiment 1a, each with a normal condition, a plausible error condition, an implausible error condition, and a very implausible error condition, for a total of twelve conditions. Participants were presented with only one of the conditions. The cover stories were modeled on those from experiments 1a - 1d: the normal and plausible error conditions were the same, and two new error conditions were developed.

Like the Siberian grebe vignettes from experiment 2a, all three cover story types focused on perceptual error possibilities.

As in Experiment 2a, the normal cases serves as controls, and no unrealized possibility of error is mentioned; the plausible error cases mention the possibility of a plausible source of error; the implausible error cases mention the possibility of error that is subjectively less likely; and the very implausible error cases mention the possibility of an error that while logically possible, if subjectively very unlikely. As in 2a, the very implausible error cases were intended to resemble global skeptical hypotheses, by involving evil demons or alien super-psychologists to motivate the error possibilities. As in Experiment 2a, each participant only saw one case, and after reading the story they were asked to rate their agreement to the following statement "John knows the (table/animal/plant) is (red/a jaguar/a gorse bush)" on a six point likert scale where 1 = strongly disagree and 6 = strongly agree. On a subsequent page, with the vignette no longer visible, participants were then asked to rate how plausible they thought the scenario was on a six point likert scale where 1 = very implausible and 6 = very plausible. The vignettes are presented, along with mean knowledge attribution rates and plausibility assessments in the Table 11, below. The results are visualized in Figures 5, 6, and 7.

Results: Experiment 2b

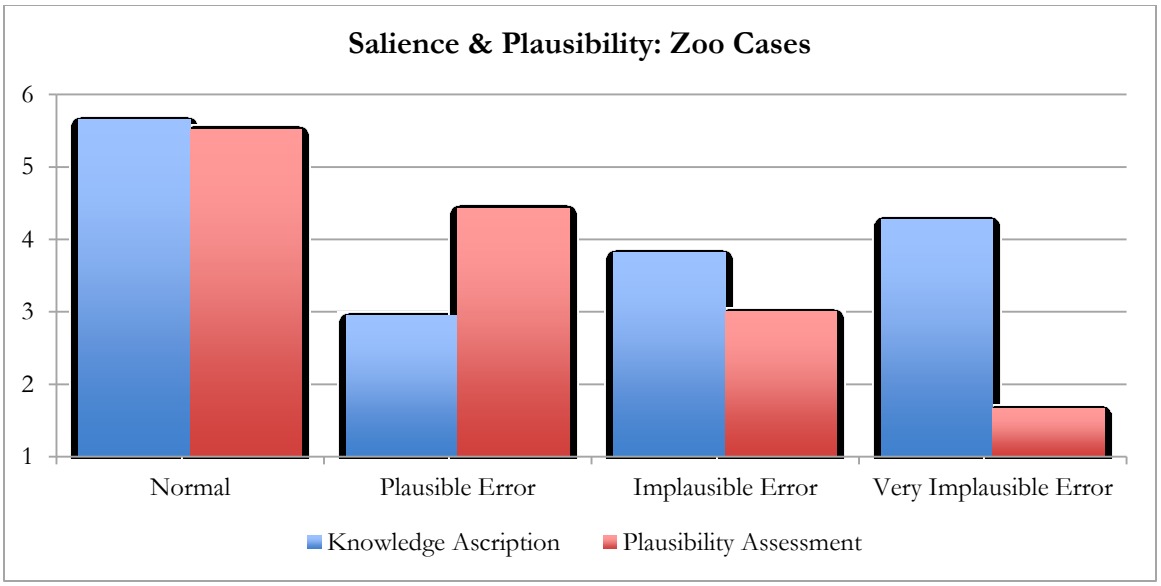


Figure 5: Experiment 2b: Zoo cases plausibility and knowledge attribution levels.

	Zoo Condition (1b)	Arboretum Condition (1c)	Furniture Condition (1c)
Normal Case	John and his friend go to the zoo. As they walk around, they pause in front of an exhibit marked "Brazilian Jaguar Enclosure". John and his friend read about jaguars from the sign, and look out and see a jaguar sleeping on the branch of a tree in the enclosure. Knowledge M = 5.65 Plausibility M = 5.52	It is a beautiful Saturday early in spring, and John goes to the local arboretum to see the plants flowering and to enjoy the weather. As he walks around, he pauses in front of a yellow flowering bush. The sign next to the bush is labeled "Gorse." In his guidebook he reads that gorse is native to Scotland, and that it is a genus of flowering plants in the family Fabaceae. Knowledge M = 4.64 Plausibility M = 5.49	John A. Doe is in a furniture store. He is looking at a bright red table under normal lighting conditions. He believes the table is red. Knowledge M = 5.18 Plausibility M = 5.50
Plausible Error Case	Normal Case + However, African leopards look very much like Brazilian jaguars, and the signs in the zoo have recently been replaced by an inexperienced crew of workers. If a zoo official had accidentally switched the signs on their exhibits, John wouldn't have been able to tell the difference between the jaguar he did see and a leopard. Knowledge M = 2.95 Plausibility M = 4.43	Normal Case + However, nearby in the arboretum there is an exhibit of an Asian forsythia which looks remarkably similar to gorse. If the arboretum staff had mixed up the signs, John wouldn't have been able to tell the difference between the gorse bush he did see and a forsythia bush. Knowledge M = 3.43 Plausibility M = 4.88	Normal Case + However, a white table under red lighting would look exactly the same to him, and he has not checked whether the lighting is normal, or whether there might be a red spotlight shining on the table. Knowledge M = 2.97 Plausibility M = 2.36
Implausible Error Case	Normal Case + However, the American mountain lion is about the same size as a jaguar, even though it doesn't have spots. If zoo officials had cleverly painted a mountain lion to look like a jaguar, John wouldn't have been able to tell difference between the jaguar he did see and a cleverly painted mountain lion. Knowledge M = 3.82 Plausibility M = 3.00	Normal Case + However, it's possible to make plastic plants that look very realistic. If the arboretum staff had put a plastic replica in the ground, John wouldn't have been able to tell the difference between the gorse bush he did see and a plastic replica. Knowledge M = 3.87 Plausibility M = 4.46	Normal Case + However, if the furniture store accidentally focused magenta and yellow spotlights onto a white table, it would look exactly the same to him. He has not checked whether the lighting is normal, or whether there might be magenta and yellow spotlights shining on the table. Knowledge M = 3.84 Plausibility M = 2.79
Very Implausible Error Case	Normal Case + However, a malevolent demon could produce the illusion of a jaguar even if there wasn't actually one in the enclosure. If a malevolent demon had produced the illusion of a jaguar, John wouldn't have been able to tell the difference between the jaguar he did see and an illusion. Knowledge M = 4.27 Plausibility M = 1.66	Normal Case + However, if little garden gnomes secretly lived in the arboretum, and one night they decided to dig up the gorse bush and replace it with a very similar looking forsythia bush, John wouldn't have been able to tell the difference between the gorse bush he did see and a forsythia bush. Knowledge M = 4.22 Plausibility M = 1.63	Normal Case + However, if little green elves in the ceiling of the store focused magenta and yellow spotlights onto a white table, it would look exactly the same to him. He has not checked whether the lighting is normal, or whether little green elves might be shining magenta and yellow spotlights on the table. Knowledge M = 4.3 Plausibility M = 1.53

TABLE 11: Salience and Plausibility Experiment 2b - 2d: Salience and plausibility replication.

Experiment 2b replicates all the major results of 2a, showing that the inverse relationship between the plausibility of an error possibility and salience effects is a robust one. As the plausibility of an error possibility decreases (red bars in Figure 5), knowledge attribution goes up (blue bars). I'll focus first on the zoo vignettes. Because knowledge attribution in all four zoo conditions are non-normal, as assessed by a Shapiro-Wilk Test, at the $p<.001$ level (normal case, $W=0.637$; plausible error, $W=0.865$; implausible error, $W=0.887$; very implausible error, $W=0.833$), I analyzed the results with non-parametric statistics.

An omnibus Kruskal-Wallis showed a significant effect for condition $H(3)=54.55$, $p=.000$, $\eta^2=.315$. Post-hoc tests showed significant differences between the normal zoo case ($n=48$; $M=5.65$; $SD=0.52$) and the zoo plausible error case ($n=37$; $M=2.95$; $SD=1.67$), $H(1)=74.48$, $p<.000$, $\eta^2=.529$ (a "large" effect); as before, the result is consistent with past experimental demonstrations of salience effects. Moreover, it shows that when an unrealized error possibility is considered highly plausible, as in the zoo plausible error case ($M=5.51$; $SD=0.95$) it can push intuitive knowledge attributions well below the midline of the likert scale.

The most important comparisons, though, are between the normal zoo case and the zoo very implausible error case, and between the zoo plausible error case and the zoo very implausible error case. In the zoo vignettes, unlike the Siberian grebe vignettes of experiment 2a, there was a significant difference between knowledge attribution in the normal zoo case and the zoo very implausible error case ($n=44$; $M=4.27$; $SD=1.78$), $H(1)=39.83$, $p<.001$, $\eta^2=0.189$ (conventionally a "small" effect). However, this shouldn't been seen as showing that plausibility has doesn't influence the size of salience effects. First, knowledge attribution in the normal zoo case was near ceiling, and in so far as mentioning an implausible error possibility reduces intuitive knowledge attribution among some participants, it creates the potential for some kind of salience effect. In this case the effect-size was small. Moreover, knowledge attribution in the zoo very implausible error case was significantly higher than in the zoo plausible error case, $H(1)=-34.56$, $p=.009$, $\eta^2=0.124$. So, while subjectively implausible defeaters *can* induce salience effects, their relative effect size of implausible defeaters ($\eta^2=0.189$) is modest by comparison the effect size of plausible defeaters ($\eta^2=.529$).

To directly test the relationship between plausibility assessments and knowledge attributions, I conducted a regression of knowledge and plausibility assessments. The results confirmed the findings of experiment 2a. The regression model used plausibility assessments as a predictor for knowledge attribution in the three cases that mentioned an unrealized possibility of error. There was a significant inverse relationship between plausibility rating and knowledge attribution at the p <.006 level; see Table 12.[100]

	B	SE *B*	Beta
Constant	4.36	.296	
Plausibility rating	-0.232	0.083	-.237*

*r^2 = .056, p <.006

Table 12: Regression Model for Experiment 1b: Zoo Cases

Results: Experiment 2c.

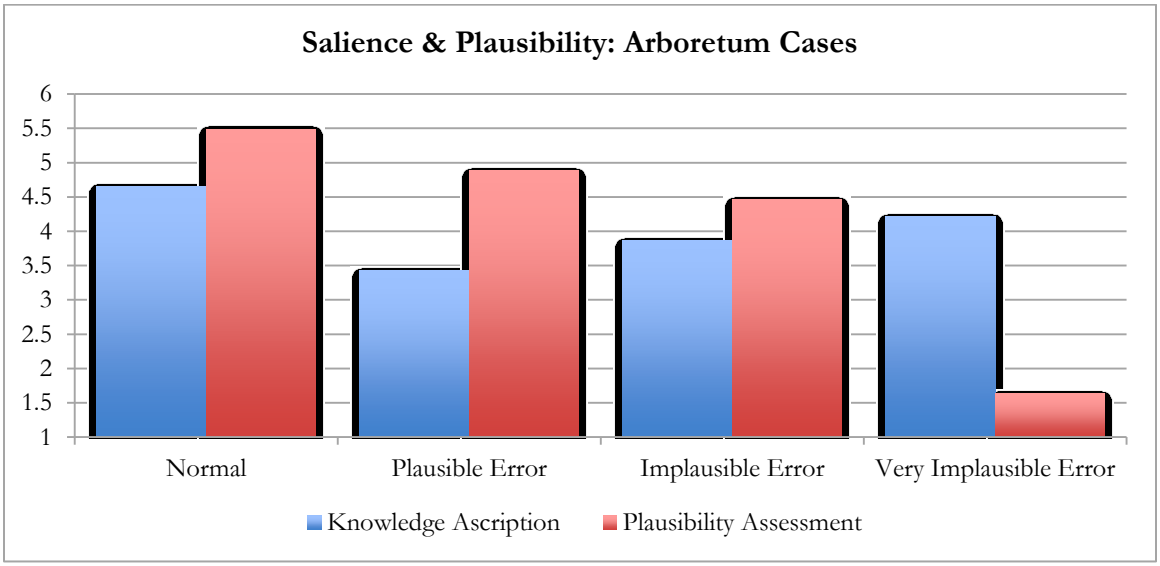


Figure 6: Experiment 2c: Arboretum cases plausibility and knowledge attribution levels.

Results from the arboretum cases in experiment 2c replicate the findings above. Because all four conditions are non-normal, as assessed by a Shapiro-Wilk Test, at the p<.004 level (normal arboretum, W=0.827; arboretum plausible error case, W=0.925;

[100] Demographic analysis: There were no sig correlations for knowledge attribution and gender (p=0.30), age (p=0.53), race (p=0.76), education (p=0.12), philosophy education (p=0.43). There were no significant correlations for plausibility attribution and gender (p=0.673), age (p=0.266), race (p=0.54), education (p=0.286), or philosophy education (p=0.40).

arboretum implausible error case, W=0.901; arboretum very implausible error case, W=0.889), I analyzed the results with non-parametric statistics.

An omnibus Kruskal-Wallis showed a significant effect for condition $H(3)=15.56$, $p=.001$, $\eta^2=.09$ Post-hoc tests showed significant differences between the normal case (n=43; M=4.65; SD=1.49) and the arboretum plausible error case (n=49; M=3.43; SD=1.51), $H(1)=42.64$, $p<.001$, $\eta^2=.1549$ (a "small" effect), as before this result is consistent with past experimental demonstrations of salience effects.

Focusing on the most important comparisons, in this cases there was no statistically significant difference between the normal arboretum case and the arboretum very implausible error case (n=41; M=4.22; SD=1.55) $H(1)=15.34$, $p=1$.[101]

To again directly test the relationship between plausibility assessments and knowledge attributions in the arboretum cases, I conducted a regression. The results confirmed the findings of previous experiments. The regression model used plausibility assessments as a predictor for knowledge attribution in the three cases that mentioned an unrealized possibility of error. There was a significant inverse relationship between plausibility rating and knowledge attribution at the p <.004 level; see Table 13.[102]

	B	SE *B*	Beta
Constant	4.50	0.274	
Plausibility rating	-0.193	0.065	-.235*

*r^2 = .055, p <.004

Table 13: Regression Model for Experiment 1c: Arboretum Cases

Results: Experiment 2d.

Turning now to the furniture cases, the results of the study are consistent with all of the previous findings, although the picture is somewhat less clear because the sample sizes in this study were the smallest, giving it much less inferential power than the others. Nevertheless, the overall picture is much the same.

[101] Give all pair-wise comparisons:

[102] Demographic analysis: There were no significant correlations for knowledge attribution and gender (p=0.30), age (p=0.53), race (p=0.76), education (p=0.12), philosophy education (p=0.43). Nor were there significant correlations for plausibility attribution and gender (p=0.673), age (p=0.266), race (p=0.54), education (p=0.286), or philosophy education (p=0.40).

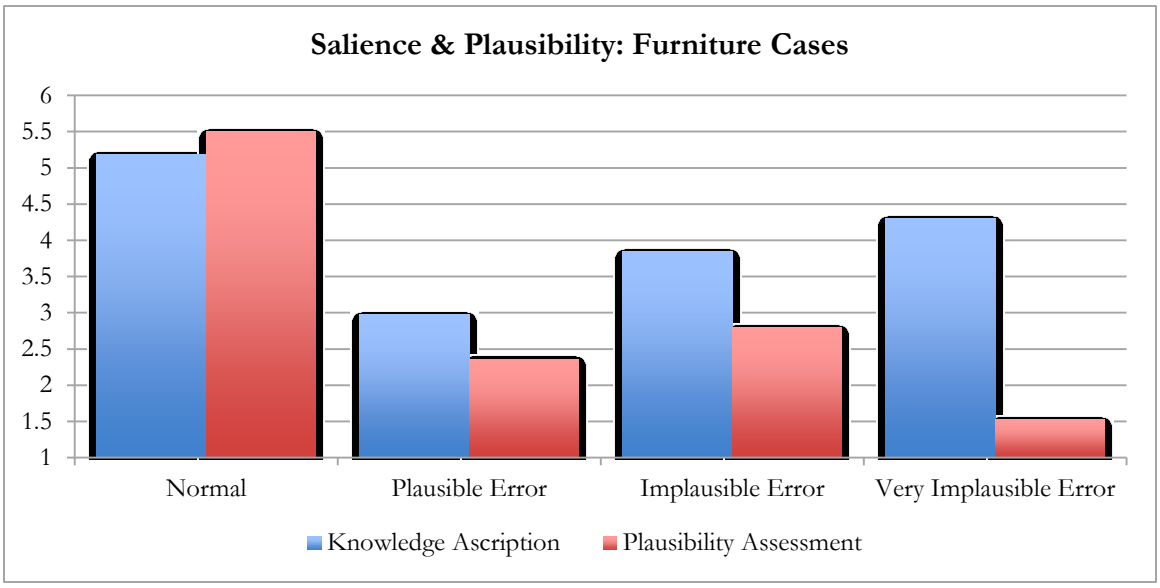


Figure 6: Experiment 2d: Furniture cases plausibility and knowledge attribution levels.

Because three of the four conditions are non-normal, as assessed by a Shapiro-Wilk Test, at the $p<.005$ level (normal furniture case, $W=0.686$; plausible furniture error case, $W=0.887$; implausible furniture error case, $W=0.920$; very implausible furniture error case, $W=0.880$), I analyzed the results with non-parametric statistics.

An omnibus Kruskal-Wallis showed a significant effect for condition $H(3)=26.07$, $p<.000$, $\eta^2=.253$. Post-hoc tests showed significant differences between the normal furniture case ($n=22$; $M=5.18$; $SD=1.26$) and the plausible furniture error case ($n=33$; $M=2.97$; $SD=1.43$), $H(1)=40.39$, $p<.000$, $\eta^2=.425$ (a "large" effect), once again this result is consistent with past experimental demonstrations of salience effects.

Focusing on the most important comparison for understanding the role of plausibility judgments, in this cases there was no statistically significant difference between the normal furniture case and the very implausible furniture error case ($n=30$; $M=4.30$; $SD=1.60$) $H(1)=16.63$, $p=.272$, $\eta^2=.095$. As in experiments 2a, 2b, and 2c, there is a trend toward lower knowledge attribution in all four very implausible error cases, even if it is not at a statistically significant level. But, as before, the most important comparison is between effect sizes. In this case, introducing a subjectively plausible possibility of error produces salience effects that are four times larger than the subjectively implausible error cases.

Finally, to directly test the relationship between plausibility assessments and knowledge attributions in the arboretum cases, I conducted a regression analysis. The results are consistent with the findings of previous experiments. The regression model used plausibility assessments as a predictor for knowledge attribution in the three cases that mentioned an unrealized possibility of error. There was a marginally significant inverse relationship between plausibility rating and knowledge attribution at the $p<.084$ level; see Table 14.[103] These results don't reach the critical value of .05 largely because the sample size was too small. Nevertheless, the results are entirely consistent with 2a, 2b, and 2c.

	B	SE *B*	Beta
Constant	4.13	0.32	
Plausibility rating	-0.22	0.12	-.190*

*$r^2 = .037$, $p < .084$

Table 14: Regression Model for Experiment 1d: Furniture Cases

General Discussion:

These four experiments reveal an important overall pattern: the subjective plausibility of an unrealized possibility of error plays a sizeable roll in determining its salience. When error possibilities are plausible, salience effects are large. As they become less plausible, salience effects shrink. When error possibilities become truly implausible, salience effects become insignificant or disappear entirely. In three of the four experiments, there was no statistically significant difference between the normal case, where no error possibility is mentioned, and the very implausible error case, where a subjectively very implausible possibility of error is mentioned. While there is an overall trend toward lower knowledge attribution in very implausible error cases in all four story types, the size of the difference was very small, which is why it rose to the level of significance only once. Comparing effect sizes here is the most useful way to understanding the role of plausibility in generating salience effects. In comparing the normal cases to the plausible error cases, the effect sizes were, Siberian grebe: $\eta^2=.429$; Zoo: $\eta^2=.529$; Arboretum: $\eta^2=.154$; Furniture: $\eta^2=.425$. All of which are large effects, with the exception of the arboretum, which is small. By contrast,

[103] Demographic analysis: There were no significant correlations for knowledge attribution and gender (p=0.958), age (p=0.289), race (p=0.182), education (p=0.510), philosophy education (p=0.225). Neither were there significant correlations for plausibility attribution and gender (p=0.813), age (p=0.178), race (p=0.082), education (p=0.990), or philosophy education (p=0.770).

when we compare the normal case and the very implausible error case, the effect sizes were small to insignificant, Siberian grebe: η^2=.006; Zoo: η^2=.189; Arboretum: η^2=.025; Furniture: η^2=.095. The one exception here is the Zoo case, which would be classified as a small effect, but it is nevertheless close to a third of the size of the standard Zoo case comparison.

As I will argue in just a moment, these results undermine the diagnoses of the skeptical problem offered by linguistic contextualists and invariantists who rely upon an indifference to the plausibility of error. But before pressing the point, it's worth addressing exactly how similar the very implausible error cases used in these experiments are to the kinds of skeptical hypotheses that can be found in discussions of radical skepticism. Indeed, it could be objected that these cases not analogous to the skeptical hypotheses typical of AI. Are these error cases like radical skepticism? There are important similarities and differences. Both focus on perceptual belief and perceptual error. In both the basis of error is not *unknown* to the culture, people know what demons, aliens, and gnomes are supposed to be, even if they've never encountered one. A major difference is that these cases are not global, though, and in this way are not *radical*. But should this be a reason to think they'd be *less* effective at generating salience? One reason to think not is that they may be somewhat more plausible than the kinds of error at the route of global skeptical hypotheses. Indeed, how truly implausible are these hypotheses? How implausible was Descartes's demon to his contemporaries? It would be problematic for the interpretation I pursue here if the scenarios I presented to the participants were much more implausible than those typically used to advance arguments like the argument from ignorance.

Perhaps the very implausible error cases are not as implausible as one might first imagine. In a recent paper exploring the American public's persistent fascination with conspiracy theories in the realm of politics, J. Eric Oliver and Thomas J. Wood found that more than 50% believed in at least one conspiracy theory, broadly construed, including, for example, that the 9/11 attacks were masterminded by American politicians to justify a war for oil, that the CIA assassinated John F. Kennedy, or that some secret organization like the Masons truly runs the world. In diagnosing what Hofstadter called "the paranoid style of American politics" Oliver and Wood found that susceptibility to "conspiricism" was predicted by the tendency toward "magical thinking" - the tendency to explain phenomena

as the effect of unobservable causal forces like angels and demons[104] - and the tendency towards "manicheanism" - the belief that the world is a battleground over which the forces of good and evil fight. Using results from Harvard University's large Cooperative Congressional Election Study (CCES),[105] they found that a third of all Americans believe in ghosts or that some people have ESP, moreover, they found that 60% of Americans believe in the devil and in angels.[106] A recent study commissioned by the National Geographic Society found that a similar number of Americans believe that there is evidence that UFOs have visited earth.[107] Given this widespread acceptance of the actors behind the possibility of error in the very implausible error cases I deployed in these experiments, it is defensible to conclude that basing them on elves, aliens, demons, and gnomes does not unduly bias the tests because they are *more* implausible than those that are the stock and trade of epistemology classes. But if many people actually believe in these kinds of agents, how implausible are they really? It turns out that even among people prone to conspiracy theories, despite the fact that they will sometimes invoke supernatural agents in explanations, they rarely invoke them regularly.

To sum up then, the cases considered here experimentally are not different from traditional skeptical hypotheses in the plausibility of their source of error. Both traditional hypotheses and these invoke sources of error that are familiar: brains in vats are the stuff of movies, and demons the stuff of religion. Compare this to the simulation hypotheses at the heart of Nick Bostrom's global skeptical argument.[108] Bostrom considers the possibility that we *are* a computer simulation -- not merely wet brains plugged into a computer, but that we are in fact entirely the product of a computer. This kind of hypotheses is not just subjectively implausible, it's quite counter-intuitive. The major distinction, then, is that radical skepticism generally is global, and targets all empirical beliefs, whereas these scenarios target only specific beliefs. Leaving aside that many philosophers, including those that seem to endorse indifference to plausibility, have been quite happy with the analogy, the fact that these local implausible skeptical arguments are not global, might make them more subjectively plausible. It's a hard call, I admit, but what is more plausible, that a malevolent

[104] R. Shweder, 1977.
[105] Stable URL: http://projects.iq.harvard.edu/cces/home.
[106] J. Oliver & T. Wood, 2012: 21.
[107] Stable URL: http://goo.gl/FnNxNf (retrieved June 6th, 2014).
[108] N. Bostrom, 2003.

demon has decided to deceive you in your trip to the zoo, or that he has decided to deceive you about every one of your beliefs? Considerations can be raised on both sides. A virtue of the scenarios tested here is that, were the error possibilities realized, they would not render every belief the agent had false, preserving the rationality of relying on perception generally. But perhaps *if* an implausible scenario *were* in the offing, it makes more sense to think that a demon or alien intelligence would want to deceive us about everything rather than one thing because we are the subject of some kind of experiment. I don't think these considerations are decisive in either direction, while there is no clear case that the scenarios considered here are *more* plausible than global skeptical hypotheses, neither is there a good case to be made that they are distinctly *less* plausible. Consequently, they stand as good tests of our intuitive responses to the kinds of sources of error used to motivate radical skepticism.

How committed are diagnoses of skepticism to indifference to plausibility? It would seem that those offered by invariantists like Williamson and Hawthorne are deeply committed, but what then of linguistic contextualists? They surely do not need to admit, as we saw in section 4, that any mention of a skeptical hypothesis is sufficient to induce skepticism. But it is a delicate issue; to the extent that they aim at a genuine diagnosis of how skeptical arguments get a grip on us, they must have something to say about the circumstances under which skeptical hypotheses work, and when they don't. And this, it is critical to mention, follows from the diagnostic nature of their endeavor. Contextualism does not aim to answer the normative question of when skeptical hypotheses *deserve* our attention, but only the empirical question of how they do *capture* it. Stewart Cohen carefully follows Lewis here, but does not offer much detail, observing that,

> How precisely do the standards for these predicates get determined in a particular context of ascription? This is a very difficult question to answer. But we can say this much. The standards are determined by some complicated function of speaker intentions, listener expectations, presuppositions of the conversation, salience relations, etc.—by what David Lewis calls the conversational score.[109]

DeRose adopts what he calls a "skeptic-friendly" formulation of the argument from ignorance. He acknowledges that the skeptic is not always successful at raising the

[109] S. Cohen, 1999: 61.

standards, and merely mentioning a defeater doesn't always result in the upward pressure on knowledge. It's worth quoting him at length to see how he attempts to navigate this issue,

> [I]n presenting the contextualist strategy I have above assumed a skeptic-friendly version of contextualism - one according to which the philosophical skeptic can (fairly easily), and does, succeed in raising the standards for knowledge in such a way as to make her denials of knowledge true. Some contextualists may think that it's not so easy to so raise the standards for knowledge, and that a determined opponent of the skeptic can, by not letting the skeptic get away with raising them, keep the standards low. But the important point is to identify the mechanism by which the skeptic at least threatens to raise the standards for knowledge. Whether the skeptic actually succeeds against a determined opponent in so raising he standards is of little importance. To safeguard ordinary claims to know while at the same time explaining the persuasiveness of the skeptical arguments (which is the goal of his strategy), the contextualist can provisionally assume a skeptic-friendly version of contextualism, leaving it as an open question whether and under which conditions the skeptic actually succeeds at raising the standards. The contextualist's ultimate point will then be this: To the extent that the skeptic does succeed, she does so only by raising the standards for knowledge, and so the success of her argument has no tendency to show that our ordinary claims to know are in any way defective.[110]

By my lights, DeRose's position here is more than a little puzzling. If his goal is a genuine *diagnosis*, then "leaving it open…*under what circumstances* the skeptic succeeds" falls far short of a full account. It's worth noting here that DeRose and other contextualists focus on the skeptic's ability to *raise standards*. Simply assuming that the skeptic is successful passes over a significant part of diagnosis – i.e., understanding how the skeptic overcomes our ordinary indifference to implausible error possibilities. How far short depends on how skepticism works. Suppose that linguistic contextualism is true, and that it is very easy to overcome indifference to implausibility: in this case much of the diagnosis of skepticism will then be the recognition that the skeptic takes advantage of changes in context. However, if we again assume that linguistic contextualism is true, but that it is very hard to overcome indifference to implausibility, then much of the diagnoses will remain a mystery. Diagnosis aims at understanding, and insofar as we've seen here that ordinary knowledge attribution is indifferent to implausible defeaters, we still have much to understand in how skeptics

[110] K. DeRose, 1995: 6.

succeed in pressing their case.[111] But of course DeRose's true goal is to safeguard ordinary knowledge, and this is why being skeptic-friendly suites his dialectical purposes. But if understanding skepticism is the goal, and understanding skepticism is a way of understanding the nature of our epistemic psychology -- and ultimately ourselves -- then we should aim at accurately describing the problem, rather than preserving a desired solution.

6. Repetition and Cartesian Priming:

If it is part of our ordinary epistemic practice to ignore implausible error possibilities, how does the skeptic get us into the game? I suspect that it is very complicated. Many diagnostic approaches to the skeptical problem assume skeptical arguments are intuitive. The position I would take is that *some* skeptical arguments are intuitive to *some* people, in *some* contexts. That is to say, there is a complicated interaction between the argument on offer, the individual assessing it, and the status of the context they inhabit. We just saw that features of the hypothesis matter. In the next chapter I'll explore the role of individual differences in assessing skeptical arguments. Here I want to look at context manipulations.

I believe one powerful tool the skeptic has for manipulating the context is priming, and that Descartes may have been one of the first to discover this, however inadvertently. In Meditation I, Descartes does not introduce the idea of demon deception immediately. Rather, he works up to it by first presenting cases of perceptual error, then madness, then dreaming, and then, only once he has accustomed his audience to the idea that there are types of perceptual error that are difficult or impossible to distinguish from veridical perception, does he unveil demon deception. I believe that this sequential presentation of error possibilities, from the most familiar and plausible, to the most unfamiliar and implausible, tends to overcome our ordinary indifference to implausibility. In essence, Descartes primes his readers with a tacit kind of slippery slope argument, where at each step of the way the reader is invited to notice that a plausible kind of error seems to undermine knowledge, and then is invited to notice that another kind of error, though less plausible, would undermine knowledge in a similar way. By working up to the global skeptical hypotheses by easy stages, I believe Descartes is able overcome our ordinary indifference to

[111] I should make clear here that while I will offer a contextualist theory of knowledge attribution, it is not a linguistic theory. In the next chapter I will offer a version of psychological contextualism. The salience effects I've demonstrated here are better explained as the result of two different strategies for assessing knowledge, and not as the product of the underlying semantics of "know."

the implausible, because each new hypotheses tends to be compared to the most recently considered one. One way, then, to see how the heuristic rejection of implausible hypotheses works is that it is not based on absolute plausibility judgments, but on comparative judgments. In a typical context with typical standards, a global skeptical hypothesis is just too implausible to warrant consideration, and so we ignore it, and we do not penalize those who ignore it for failing to rule it out. But, the distance can be closed by introducing error possibilities that are not too distant to be ruled out. Once the context includes an error possibility, the boundary of relevance expands by comparison, and the stage is set to take the next incremental step. In this way the skeptic can overcome our ordinary indifference to implausible error possibilities by taking advantage of another epistemic heuristic, one that aims at consistency, whereby we aim to treat similar cases the same way. Something like this heuristic seems to be at work in typical sorites paradoxes, from the heap to Theseus's ship, and all the slippery slope arguments in between. Used locally a consistency heuristic is useful, and probably reasonably accurate. Used repeatedly it can lead to familiar paradoxes. The sequential presentation of error cases can take advantage of this as a way of overcoming our ordinary indifference to implausible error possibilities.[112]

6.1 Experiment 3a

Methods:

To test whether priming with increasingly implausible skeptical hypotheses affected salience, 66 participants from the United States (age *M*=33; gender: 38% female; education: 60% bachelors degree or above) were recruited online from Amazon Mechanical Turk and compensated for their participation at minimum wage. Participants belonged to a single condition, and read all four cases from experiment 2a in the following sequence: normal case, plausible error case, implausible error case, and very implausible error case. After reading each case participants were asked to rate their agreement to the following statement "John knows the bird is a Siberian grebe" on a six point likert scale where 1 = strongly

[112] Incidentally, it is worth noting that many textbook introductions to skepticism seem to adopt the same sequential, incremental introduction to radical skeptical hypotheses. See for instance M. Devitt, 1991: 62, and A. Hazlett, 2014: 71 - 76. Swain, Alexander, and Weinberg (2008) found evidence of order effects in an empirical study of intuitions about epistemological thought experiments, particularly K. Lehrer's "true-temp" cases. They found that when presented with a clear case of unfounded belief first (a psychic's guess), people tended to attribute knowledge in true-temp cases a higher level than when first presented with a case of well-grounded belief (a scientist's study). These differences in responses to the same true-temp cases suggest contrastive evaluation (as opposed to consistency).

disagree and 6 = strongly agree. On a page following the normal case, with the vignette no longer visible, participants answered a question designed to check their attention: six participants were eliminated for failing this question. On a further page, with the vignette no longer visible, participants were then asked to rate how plausible they thought the scenario was on a six point likert scale where 1 = very implausible and 6 = very plausible. Participants were then asked to fill in a standard demographic questionnaire. The vignettes are presented, along with mean knowledge attribution rates and plausibility assessments in Table 15 below. The results are visualized in Figure 8 below.

	Independent Presentation Mean Knowledge Ascription	Sequential Presentation Mean Knowledge Ascription
Normal Case	4.50	4.85
Plausible Error Case	2.15	2.37
Implausible Error Case	3.39*	2.72*
Very Implausible Error Case	4.11**	3.18**

TABLE 15: Experiment 3a: Salience and Plausibility Sequential Presentation. Knowledge ascription means indicate participants response to the following prompt: "Please indicate the extent to which you agree or disagree with the following statement: 'John knows the bird is a Siberian grebe'." Responses were measured on a six point likert scale where 1 = strongly disagree and 6 = strongly agree. Significant differences between conditions is marked by * at the $p<.05$ level, and ** at the $p<.000$ level; Significant differences between cases within a condition are marked by column position: knowledge ascriptions in different column positions are significantly different at the $p<.000$ level.

Experiment 3a Results:

Figure 8 presents a clear picture of the power of priming in overcoming ordinary indifference to implausible error possibilities. When a very implausible case is encountered on its own, participants treat it almost exactly like a case where no error has been mentioned. After sequential priming, however, a large proportion of participants treat it like a plausible salience case.

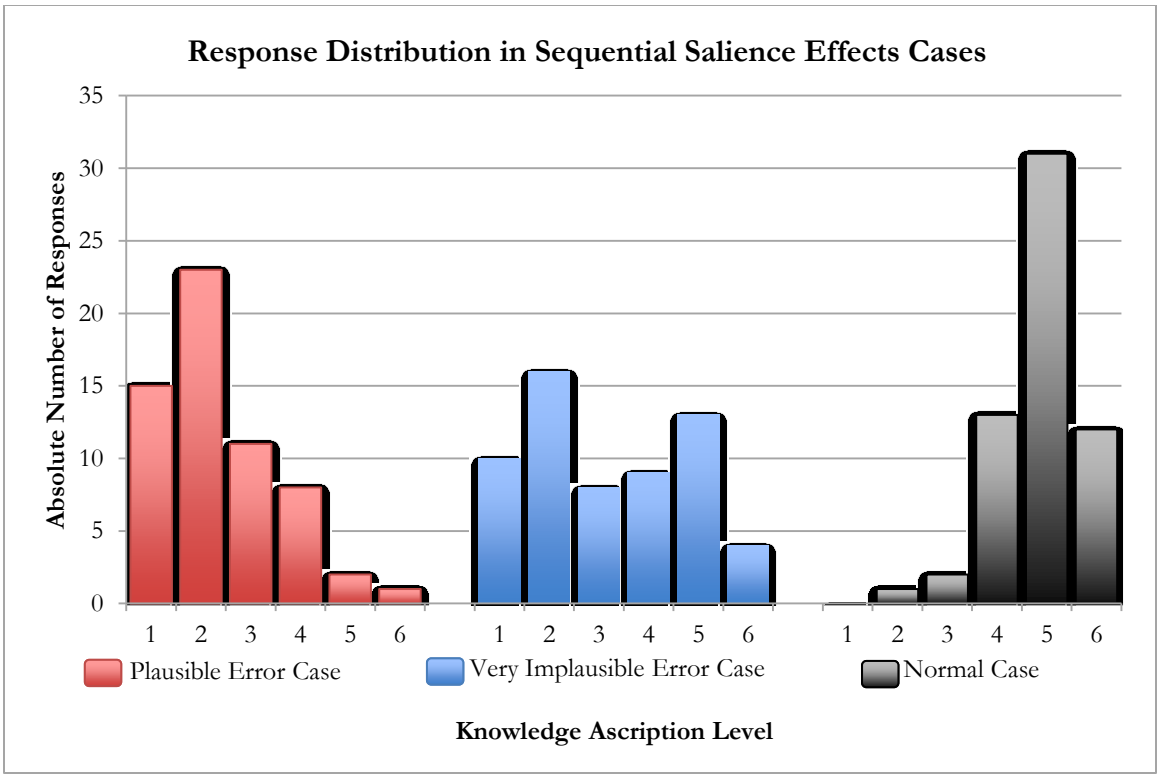


Figure 8. Experiment 3a: Response distributions in sequentially presented salience effects cases illustrate the defeat structure of skeptical cases. Column counts indicate the number of participants at that level in response to the following prompt: "Please indicate the extent to which you agree or disagree with the following statement: 'John knows the bird is a Siberian grebe'." Responses were measured on a six point likert scale where 1 = strongly disagree and 6 = strongly agree. The modal distribution in the Normal case is "5", while in the Plausible Error case it is "2". In the Very Implausible Error case, there are two modes, the first at "2" is below the midline, and the second at "5" is above the midline.

Looking first at plausibility judgments within the sequential presentation can help establish how participants viewed the cases. Visual inspection suggested that the distribution of plausibility assessments was non-normal (normal case, M=5.62; SD=0.614; plausible error, M=5.70; SD=0.591; implausible error, M=3.02; SD=1.41; very implausible error, M=1.35; SD=1.07). This was confirmed by a Shapiro-Wilk Test, which was significant for plausibility assessment in all four cases in the sequential condition at the p<.000 level (normal, W=.626, p<.000; plausible error, W=.553, p<.000; implausible error, W=.898, p=.000; very implausible error, W=.370, p<.000). Consequently, all results were analyzed using non-parametric statistics. An omnibus Kruskal-Wallis test showed a significant effect for plausibility judgment by case $H(3)$=177.39, p<.000. The normal case and the plausible error case were not significantly different $H(1)$=-5.83, p=1. The plausible error case was judged to be more plausible than the implausible error case, and the difference was

statistically significant $H(1)=86.68$, $p<.000$. The implausible error case was judged to be more plausible than the very implausible error case, and the difference was statistically significant $H(1)=50.54$, $p<.000$. This shows that the ordering was appropriate, and consistent with the independent presentations in experiments 2a - 2d. Plausibility assessments in the plausible error case were significantly above the midline ($t(59)=70.95$, $p<.000$), and they were significantly below the midline in the very implausible error case ($t(59)=9.77$, $p<.000$). This shows that participants did perceive the plausible error case as plausible, and that they did perceive our skeptical stand-in, the very implausible error case, as genuinely implausible.

Looking next at knowledge attribution within the sequential presentation, visual inspection suggested the results were not normally distributed. Because a Shapiro-Wilk Test was significant for knowledge attribution in all four cases in the sequential condition at the $p<.000$ level (normal, $W=.843$, $p<.000$; plausible error, $W=.874$, $p<.000$; implausible error, $W=.911$, $p=.000$; very implausible error, $W=.900$, $p<.000$) the results were analyzed using non-parametric statistics.

An omnibus Kruskal-Wallis test showed a significant effect for knowledge attribution by case $H(3)=86.7$, $p<.000$, $\eta^2=.363$. Post-hoc tests showed significant differences in knowledge attribution between the normal case ($n=60$; $M=4.85$; $SD=0.84$) and the plausible error case ($n=60$; $M=2.37$; $SD=1.19$) $H(1)=107$, $p<.000$, $\eta^2=.592$ (a "large" effect). This result is consistent with experiments 2a - 2d. There was also a statistically significant difference between the normal case and the implausible error case ($n=60$; $M=2.72$; $SD=1.29$) $H(1)=91.52$, $p<.000$, $\eta^2=.501$ (also a "large" effect). Most importantly, there was a statistically significant difference between the normal case and the very implausible error case ($n=60$; $M=3.18$; $SD=1.60$), which is the stand-in for a global skeptical hypothesis, $H(1)=71.68$, $p<.000$, $\eta^2=.271$ (a "medium" effect).

This result deserves a special emphasis: the key result of studies 2a - 2d was that there was no statistically significant difference in knowledge attribution between the normal case and the very implausible case, which is our stand-in for the radical skeptical hypothesis. This seemed to show that while salience effects are real, something more is needed to explain how the skeptic generates their threat. The present result points to an answer: a sequential presentation of increasingly implausible error possibilities can prime (some) people to experience salience effects in the skeptical case.

As before, the modal distribution of responses matters for our interpretation. Introducing an unrealized error possibility doesn't lower all participants' knowledge ascription levels, rather, it seems to work by prompting some people to fully deny knowledge. Figure 8 illustrates the point. Though left-skewed, the responses of those attributing knowledge in the normal case are normally distributed around the modal response of "5"; likewise, though right-skewed, the responses of those denying knowledge in the plausible error case are clustered around the modal response of "2". In the very implausible case there are two modes, one at "5" and the other at "2". This would seem to suggest that when implausible unrealized error possibilities are introduced after sequential priming, some people respond in a fully skeptical way, and other people respond in a fully anti-skeptical way. Some people feel the pull of skepticism, and some do not. When they do, they feel it fully. When they don't, they don't feel it at all.

Finally, a case-by-case comparison of knowledge attribution in the sequential presentation in this experiment and the independent presentation condition in experiment 2a further supports the conclusion that priming overcomes our ordinary indifference to implausible error possibilities. Knowledge attribution in the independent normal case (n=40, M=4.5) did not significantly differ from the sequential normal case (n=60, M=4.85) U=1,411, p=.109. Knowledge attribution in the independent plausible error case (n=52, M=2.15) did not significantly differ from the sequential plausible error case (n=60, M=2.37) U=1,810, p=.128. Knowledge attribution in the independent implausible error case (n=54, M=3.39) was significantly higher than in the sequential implausible error case (n=60, M=2.72) U=1,276, p=.046, r=-0.185 (a "small" effect size). Crucially, knowledge attribution in the independent very implausible error case (n=56, M=4.11) was significantly higher than in the sequential very implausible error case (n=60, M=3.18) U=1,152, p=.003, r=-0.276 (a "medium" effect size). Once again, this shows when people are appropriately primed, skeptical hypotheses appear more epistemically threatening than when those hypotheses are presented individually. Figure 9 visualizes the comparisons.

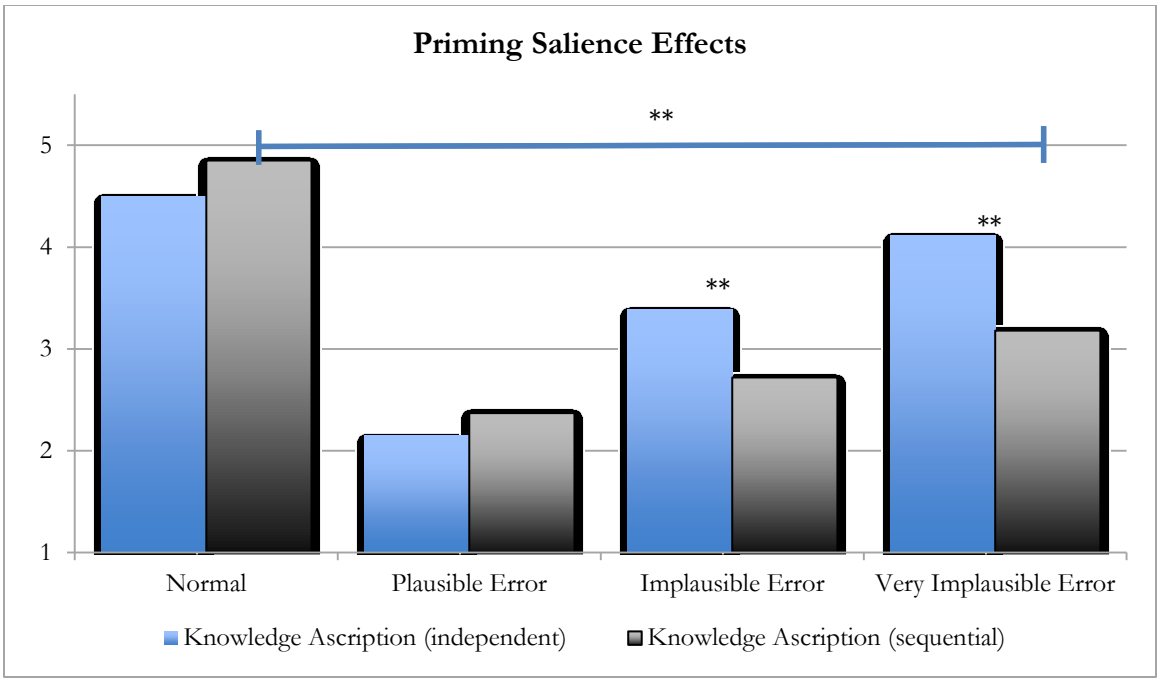


Figure 9: Experiments 2a and 3a: A comparison of mean knowledge ascription when salience cases are presented independently, and when they are presented sequentially. Comparisons marked by ** are significant at the *p*<.000 level.

One possible explanation of this pattern is that the repeated presentation of unrealized possibilities of error makes the errors appear *subjectively more* plausible to participants. In other words, on this hypothesis, the priming works by encouraging participants to see error possibilities as more plausible than they ordinarily would, and then they go about attributing knowledge in the same manner. This does not appear to be the case in this experiment: because participants tended to judge the cases as *less* plausible than when the very same stories are presented independently.[113]

Discussion:

The results of this study are support for the idea that what I call Cartesian priming, which is sequentially presenting increasingly implausible error possibilities, is effective at

[113] Plausibility assessments in the independent normal case (n=40, M=4.90) were significantly different from the sequential normal case (n=60, M=5.62) U=698, p<.001, r=-0.001. Plausibility assessments in the independent plausible error case (n=40, M=5.37) were significantly different from the sequential plausible error case (n=60, M=5.70), U=1279, p<.044, r=-0.190. Plausibility assessments in the implausible error case (n=56, M=3.8) were significantly different from the sequential implausible error case (n=60, M=3.02) U=1193, p<.006, r=-0.255. Plausibility assessments in the very implausible error case (n=56, M=1.71) were significantly higher from the sequential very implausible error case (n=60, M=1.35) U=1420, p<.049 r = -0.183. This suggests that plausibility is not the route to overcoming indifference to implausible error possibilities.

overcoming our ordinary indifference to implausible error possibilities. To review the primary results of this study, the same cases that are too implausible to generate salience effects when presented on their own (the "very implausible error cases" from experiments 2a - 2d) can generate large salience effects when people have been primed for them (experiment 3a). Moreover, looking just at participants' reactions to the very same stimulus material (the "very implausible error case" from the Siberian grebe/Gadwell duck stories), we find that knowledge attribution is much lower in the primed case (m=3.18) than in the independently presented case (m=4.11), which is significantly above the midline ($t(55)=18.91$, $p<.000$). We also saw that responses in the primed case are bimodal: the mean response of the major mode (M=2.05) is well below the midline, while the mean response of the minor mode (M=4.81) is well above. Skeptical hypotheses have all-or-nothing effects on those that feel their pull.

Priming, then, deserves a place in the more general diagnosis of the skeptical problem, as it can drive people from attributing knowledge even when an implausible skeptical scenario is in play, to denying knowledge. These results suggest that philosophers and the folk have a basic difference in the epistemic outlook. Radical skepticism has been an object of fascination in philosophy for a very long time, but it plays no role at all in the everyday lives of most people. This is a problem for a simple diagnosis that claims the skeptic is so persuasive because she relies on no more than our ordinary epistemic practices. Now, the philosopher might reply that AI does appeal to our ordinary practices, but that the best explanation of the anti-skeptical outlook of ordinary people is that they don't encounter radical skeptical hypotheses everyday. Once they are exposed to such a hypothesis, they recognize its intuitive force. It's a nice story, but the results of studies 2a - 2d show that people don't think skeptical hypotheses are intuitive. Ordinary epistemic practice and ordinary philosophical practice diverge. Study 3a shows that a sophisticated sort of tacit priming is a condition on moving the folk towards the philosopher.

It is significant that the priming is tacit. The presentation of skeptical hypotheses typically does not receive much attention in discussions of the conversational dynamics at play when skeptical hypotheses are mooted.[114] This suggests to me that, in general, people are not aware that their intuitive response hinges on whether they are considering a

[114] Contrast this with the attention lavished on speaker aims in Lewis' "Scorekeeping" and DeRose's "Single Scoreboard Semantics."

hypothesis alone, or after having been "softened up." A question I'll return to at the end of the next chapter is whether, if the skeptic's methods were fully explicit, we'd endorse the results they achieve. Some variations of presentation aid performance on normative tasks, even though we're not typically aware of it, as when probabilities are represented as frequencies. Other variations of presentation can have the opposite effect, as when probabilities are represented in decimal form. Whether the skeptic tacitly illuminates a continuity with normal epistemic practice, or obscures one, will be a question of keen interest in the last section of the next chapter.

But first, these results are open to reinterpretation: perhaps only repetition matters, not the sequence of presentation. On this alternative interpretation, the hypothesis in play is that people must be exposed to skeptical hypotheses of one kind or another repeatedly before they understand them, and once they understand them, they'll respond in a way similar to normal salience cases. It's a hypothesis worthy of consideration because of the potential role understanding might play in accurately describing the skeptical problem. Here I've proposed that people have an epistemic heuristic, 'ignore the implausible', when considering error possibilities. This epistemic heuristic plays the central role in the phenomena I've called ordinary indifference to implausible errors. I believe this heuristic explains much of the folk resistance to discussions of radical skepticism: it helps explain the "Aw, come on" response. However, another explanation of the "Aw, come on" response is that the folk don't get skeptical arguments, for a variety of reasons, when they first encounter them. If this were true, it would help explain the impulse to carefully set up discussions of radical skepticism. According to the objection I'm proposing, the traditional diagnostician of skepticism says something like the following: "Look, the reason students resist the radical skeptic's argument at first is they don't 'get it;' it's complicated, it involves negations, it's unfamiliar, and so they need to see how the argument works before they really understand how intuitively plausible it is. Once they do, they no longer say 'aw shucks, this is silly'." If this is right, then the life of the skeptical problem in philosophy doesn't boil down to a difference in epistemic outlook, just one of understanding.

6.2 Experiment 3b

Methods:

To test the repetition hypothesis, I designed a new study that repeats experiment 3a, but introduces a new condition as a control. The first condition is a sequential priming condition identical to the one used in 3a. The second condition is a repetition condition composed of the four very implausible error cases from experiments 2a-2d, presented in the following order: Furniture, Zoo, Arboretum, and Siberian grebe. See tables 9 & 11 to review the prompts. The study is designed so that that in both conditions, the fourth and final case the participants evaluate is the very implausible Siberian grebe error case. This should allow a direct test of the comparative roles of repetition and sequential presentation. If repeated exposure drives saliences effects, then knowledge attribution in the final case of each condition should be approximately the same.

To test this, 103 participants from the United States (age M=30.8; gender: 38% female; education: 49% bachelors degree or above) were recruited online from Amazon Mechanical Turk and compensated for their participation at minimum wage. Seven subjects were eliminated from the analysis for failing the comprehension question. Participants were randomly assigned to either the repetition or sequential condition. After reading each case in the sequential condition participants were asked to rate their agreement to the following statement "John knows the bird is a Siberian grebe" on a six point likert scale where 1 = strongly disagree and 6 = strongly agree. After reading each case in the repetition condition, participants were asked to rate their agreement to the following statement "John knows the (table/animal/plant/bird) is (red/a jaguar/a gorse bush/a Siberian grebe)" on the same likert scale. On a page following the normal case, with the vignette no longer visible, participants answered a question designed to check their attention: seven participants were eliminated for failing this question. On the second to last page, with the vignette no longer visible, participants were then asked to rate how plausible they thought the scenarios were on a six point likert scale where 1 = very implausible and 6 = very plausible. Participants were then asked to fill in a standard demographic questionnaire.

This study is designed to answer three basic questions. First, does sequential priming really overcome our ordinary indifference to implausible error possibilities? The results from the sequential condition, which serves as a replication of 3a, will answer this question. Second, can repetitive priming do so as well? The results from the repetition condition will

answer this question. Third, is one more powerful than the other? Comparing knowledge attribution in the final case of the sequential condition and the repetition condition will answer this question. The results are presented in Table 16 below.

Sequential Condition	Mean Knowledge Attribution	Mean Plausibility Assessment
Normal Siberian Grebe Case	4.61	5.22
Plausible Siberian Grebe Error Case	2.09	5.48
Implausible Siberian Grebe Error Case	2.39	3.04
Very Implausible Error Siberian Grebe Case	3.02	1.43

Repetition Condition	Mean Knowledge Attribution	Mean Plausibility Assessment
Very Implausible Furniture Error Case	4.70	1.35
Very Implausible Zoo Error Case	5.02	1.28
Very Implausible Arboretum Error Case	4.57	1.35
Very Implausible Siberian Grebe Error Case	4.63	1.20

Table 16. Experiment 3b: Mean Knowledge attribution and plausibility assessment in sequential and repetition presentations. Cases were presented to participants in the order on the table, starting from top to bottom. The most important comparison is knowledge attribution in the final row. In response to the same vignette, participants attributed knowledge at a much lower level in the sequential condition than in the repetition condition.

Results:

Looking at the sequential priming condition first, we find that the results replicate the results of experiment 3a.

Plausibility judgments across the cases are clearly different (normal, M=5.22; plausible error, M=5.48; implausible error, M=3.04; very Implausible Error, M=1.43). A Shapiro-Wilk Test was significant for plausibility assessment in all four conditions at the p<.000 level (normal case, W=.770, p<.000; plausible error case, W=.644 p<.000; implausible error case, W=.912, p=.000; very implausible error case, W=.481, p<.000). Consequently, all comparisons were made using non-parametric statistics. Friedman's Related-Samples Two-way Analysis of Variance by Ranks shows there is a statistically significant difference in plausibility assessments $Fr(3)$=100.15, p<.000. Once again, the

mean plausibility judgment for the plausible error case was significantly above the midline ($t(1)=15.56$, $p<.000$) and the mean plausibility judgment for the very implausible error case was significantly below the midline ($t(1)=-13.66$, $p<.000$). This shows that the cases were well ordered, and importantly, that the plausible error case struck people as subjectively plausible, while the very implausible error case, our stand-in for the radical skeptic's hypothesis, struck people as subjectively implausible. Satisfied that this basic assumption about how the vignettes were perceived is correct, we can focus on the central analysis.

Turning to knowledge attribution in the sequential condition, a Shapiro-Wilk Test was significant for all four conditions at the $p<.000$ level (normal case, $W=.892$, $p<.000$; plausible error case, $W=.828$ $p<.000$; implausible error case, $W=.835$, $p=.000$; very implausible error case, $W=.863$, $p<.000$) and so the cases were compared using Friedman's Related-Samples Two-way Analysis of Variance by Ranks. The test shows there is a significant difference in knowledge attribution across cases $Fr(3)=71.6$, $p<.000$. Post-hoc tests show that knowledge attribution in the normal case ($n=46$, $M=4.61$; $SD=1.13$) is higher than knowledge attribution in the very implausible case ($n=46$, $M=3.02$; $SD=1.82$) $Fr(1)=1.11$, $p<.000$). This replicates the main finding of experiment 3a. Sequential priming lowers knowledge attribution in the very implausible case far below the level at which people attribute knowledge when they encounter the case on its own.

Looking now at the repetition condition, we find that repetitive presentation does not have any priming effect on salience in either direction. The plausibility assessments show that participants found all four cases very implausible. Because a Shapiro-Wilk Test was significant for all four conditions at the $p<.000$ level (very implausible furniture error, $W=.418$, $p<.000$; very implausible zoo error, $W=.306$, $p<.000$; very implausible arboretum error, $W=.411$, $p=.000$; very implausible grebe error, $W=.850$, $p<.000$), all results were analyzed non-parametrically.

Plausibility assessments are significantly below the midline for all four cases (furniture, $t(46)=-15.03$, $p<.000$; zoo, $t(46)=-14.67.4$, $p<.000$; arboretum, $t(46)=-15.4$, $p<.000$; Siberian grebe, $t(46)=-22.75.4$, $p<.000$) which confirms that participants found all four cases very implausible. Comparing plausibility assessment in this condition to plausibility assessments of the same cover stories presented independently in studies 2a - 2d, we find that while plausibility assessments in all cases are very close to the bottom of the scale, there is still some variation in assessment depending on whether the case is

encountered alone, or as part of a sequence. In the very implausible furniture cases we find that plausibility attribution in the independent presentation (M=1.53, SD=1.20) is higher than in the repetition presentation (M=1.35, SD=0.971). A Mann-Whitney U test shows the difference is not statistically significant, $U(90)$=727.5, p=.754.

In the very implausible zoo cases we find that plausibility attribution in the independent presentation (M=1.66, SD=1.20) is higher than in the repetition presentation (M=1.28, SD=1.03). A Mann-Whitney U test shows the difference is statistically significant, $U(90)$=900, p=.041, r=-0.211 (a "medium" effect size).

In the very implausible arboretum cases we find that plausibility attribution in the independent presentation (m=1.67, SD=1.30) is higher than in the repetition presentation (M=1.35, SD=0.948). A Mann-Whitney U test shows the difference is not statistically significant, $U(90)$=1054, p=.467.

In the very implausible Siberian grebe cases we find that plausibility attribution in the independent presentation (M=1.714, SD=1.37) is higher than in the repetition presentation (M=1.20, SD=0.687). A Mann-Whitney U test shows the difference is statistically significant, $U(96)$=1144.5, p<.025, r=-.218 (a "medium" effect). While there is a general trend towards lower plausibility ratings in the repetition condition, the difference is only statistically significant in the 2nd and 4th cases. From this we can conclude the repetition, and thus familiarity, *does not* tend to render cases more plausible, and we may tentatively conclude that it *does* render them less plausible.

The last important comparison of plausibility assessments is between the repetitive and sequential presentations of the very implausible Siberian grebe error cases. A Mann-Whitney U test showed that there was no statistically significantly difference between plausibility assessments in the sequential condition (M=1.43) and the repetition condition (M=1.20), $U(91)$=1,152; p<.127. This establishes that, independent of their effects on knowledge attribution, sequential and repetitive presentations do not have differential influences on subjective plausibility assessments. Having established this, we can turn with confidence to the comparison of knowledge attribution after these presentations.

The results are stark and dramatic. The key comparison between conditions is of participants' knowledge attribution in the very implausible Siberian grebe error case. Remember that this was the fourth case participants encountered in both conditions. In the repetition condition, knowledge attribution (n=46, M=4.63; SD=1.25) was significantly

($t(45)=6.16$, $p<.000$) above the midline, while in the sequential condition knowledge attribution ($n=46$, $M=3.02$; $SD=1.82$) was below the midline to a marginally significant degree ($t(45)=-1.78$, $p<.081$). Comparing knowledge attribution in the two conditions, a Mann-Whitney U Test shows the difference is statistically significant $U(96)=542.5$, $p<.000$, $r=-0.402$ (a "large" effect size). This shows that sequential priming drives down knowledge attribution, while repetition seems to have no effect. To confirm this finding, we need only compare the results of the repetition condition to those in experiment 2a, when the same case was presented in isolation.

In experiment 2a the mean knowledge attribution of participants ($M=4.11$) in the very implausible Siberian grebe error cases was well above the midline, indicating a clear tendency to attribute knowledge to the agent in the vignette. When we compare results from the independent condition to knowledge attribution in the repetitive priming condition ($M=4.63$), a Mann-Whitney U test finds the difference is *not* statistically significant $U(102)=1{,}079$, $p=.148$. Because knowledge attribution after repeated exposure to implausible hypotheses is no different than after encountering an implausible hypotheses for the first time, it suggests that familiarity and understanding are not an important variables in diagnosing responses to skeptical hypothesis. As a final check of our results, we can also compare knowledge attribution in the independent condition from 2a to the sequential condition in this study ($M=3.02$): a Mann-Whitney U test confirms that they are different, $U(101)=861$, $p<.004$, replicating the key result of study 3a.

Discussion:

Up to this point I have argued that it is an entrenched part of our ordinary epistemic practice to ignore very implausible error possibilities. I then argued that the skeptic overcomes this indifference to the implausible through a kind of priming of which we are not aware. Experiment 3b was designed to test a potential challenge to this account that could be advanced to on behalf of traditional diagnoses of the skeptical problem. That challenge is simple, ordinary people respond to the radical skeptic with an incredulous stare only because the skeptic's challenge is unfamiliar. According to the challenge, once ordinary people understand the structure of the skeptic's hypothesis, they'll respond to it in much the same way they respond to normal salience cases.

The results of 3b don't support this hypothesis. Sequential priming has a powerful effect on knowledge attribution (r=-402), whereas repetition has none. This result does not show repetition cannot generate salience effects with implausible hypotheses -- perhaps longer or more varied kinds of exposure would generate similar results to sequential priming. But what it does allow us to judge is comparative power. Sequential priming is clearly powerful: it can be achieved compactly, in the course of a conversation (average time to completion in study 3a was under five minutes), and is thus clearly capable of manipulating context in real time. While it seems sensible to me to think that repeated exposure can achieve the same effect, and perhaps many of us are familiar with something similar in our own classroom experiences, there is no reason to think that it has anything like the real-time power of Cartesian priming in altering contexts.

A different worry is that Cartesian priming is an artifact of accommodation: some participants realized, or made a guess about the aims of the study, and thought they were expected to respond to the very implausible error cases the same way they responded to the plausible error cases. Accommodation is a live worry in any study, but one reason to think it was not in play here was the constant level of the plausibility assessments across independent and sequential presentations. If participants were making this guess, then one would assume they'd carry that guess over to their responses to the plausibility questions. There was no corresponding trend in response to the plausibility questions: rather than rate the very implausible cases more like the plausible ones, assessments in the sequential condition did not differ from those in the independent condition.

7. Conclusion

The results of the experiments I've reported in this chapter can be summarized quickly. Experiments 1a - 1d showed that salience effects are genuine. Experiments 2a - 2d showed that plausibility matters to salience effects, and that implausible error possibilities do not generate salience effects. Experiment 3a showed that sequential priming can overcome our ordinary indifference to implausible error possibilities. Experiment 3b showed that repeated exposure to implausible error possibilities does not generate salience effects. Each of these experiments has important implications, which I'll take up in turn.

Salience Effects:

Demonstrating that salience effects are genuine and general is very important for the debate over contextualism. As we saw (much) earlier in the chapter, linguistic contextualism is still controversial, even though it is the dominant approach to explaining the skeptical problem.[115] While many philosophers agreed that the basic pattern of intuitive response existed, the early attempts to test this claim empirically with ordinary speakers failed. Those results put considerable pressure on contextualists because their diagnostic focus has always been on ordinary epistemic practice. It's instructive to consider a move that linguistic contextualists *did not* adopt in response to the early null results: they did not adopt a modified expertise defense. Contextualists could have argued, "sure, ordinary speakers don't have salience intuitions, but no matter: ordinary speakers don't get the problem of skepticism anyways. Salience intuitions are part of *philosopher's ordinary epistemic practice*, and it's philosophers that are concerned with skepticism anyways." As we saw in Chapter 1, the initial reaction to many findings in experimental philosophy was to make just such a move. The obvious reason linguistic contextualists did not retreat in this way is that they posit semantic context-sensitivity, akin to the kind exhibited by indexicals or scalar adjectives, as the mechanism that explains the shift in intuitions. By appealing to a general feature of the language, and thus a feature of everyone's concept of knowledge, they establish the centrality of skepticism to our epistemic position. But if they adopted an expertise defense, it would have been like retreating to a "high-standards" view of knowledge, one according to which the lessons of skepticism are much less clear, because it is not obvious that the "high standards" standard is the one anyone should adopt. So, to the extant that for linguistic contextualists "the best grounds for accepting contextualism come from how knowledge attributing (and knowledge denying) sentences are used in ordinary, non-philosophic talk: What ordinary speakers will count as 'knowledge' in some non-philosophical contexts they will deny is such in others"[116] then the results presented in experiments 2a - 2d square with their argument. Knowledge attributing (and knowledge denying) sentences *really* are used differently in different contexts by non-philosophers.

---

[115] Bourget and Chalmers (2013: 12 manuscript) found that 40% of working analytic philosophers categorize themselves as contextualists, 31% as invariantists, 3% as relativists, and the remainder as holding "other positions." Interestingly, in a metasurvey question, Bourget and Chalmers found that philosophers wrongly believe invariantism is the leading view among other philosophers.

[116] DeRose, 2009: 47

However, we should stop short of thinking that these results support any particular hypothesis about the mechanism(s), semantic or otherwise, that generate the effect. These results are compatible with many proposals. Norman Malcolm, for instance, advanced a forerunner of contextualism according to which "know" is ambiguous between an ordinary "weak" sense of the word and a more "stringent" philosophical sense.[117] René van Woudenberg has recently argued that "know" is multiply ambiguous, and this accounts for the various difficulties in offering an analysis of it.[118] Allen Hazlet and Patrick Rysiew have argued that "know" is not context-sensitive, but that in some cases we mistake a conversational implicature for the content expressed by the term.[119] All of these options, and many others I haven't had room to discuss, are live at this point, and these results do not tell decisively for or against any one proposal. In Chapter 5 I'll take up the task offering an interpretation of the results, and I will defend the idea that they have a psychological basis, not a linguistic one.[120]

For now it is enough to observe that none of the many competing models on offer today makes a clear prediction about the role of plausibility in knowledge attribution, despite, as we've seen, the absolutely essential role it plays in modulating knowledge attribution under the threat of error. Experiments 2a - 2d show that when error possibilities are plausible, salience effects are large, and as error possibilities become more implausible, salience effects shrink, and even disappear.

Plausibility:

Perhaps the most important implication of experiments 2a - 2d is for the kind of diagnosis we should pursue in understanding the skeptical problem. Recognizing that there are differences in how intuitive some people find skeptical arguments can set the normative project of assessing the skeptic's argument on a more productive path. In the survey of the diagnostic literature in the introduction, we saw that the leading approach to understanding the skeptical problem begin with the assumption that arguments for radical skepticism like AI are intuitive. This account of the skeptical problem is guided by two concerns. The first,

[117] N. Malcom, 1952.
[118] R. van Woudenberg, 2005.
[119] A. Hazlett, 2009; P. Rysiew, 2001.
[120] As we saw, Nagel (2010), Hawthorne (2004), and Williamson (2005) also offer psychological accounts. Chapter 5 will explore the limitations of their proposals further.

and most basic, is to explain how an argument that leads to a conclusion so deeply at odds with our ordinary understanding of the world and our knowledge of it can have had the influence and fascination that skepticism has exercised in the west since Descartes. We would not continue to worry about an argument for so deeply implausible a conclusion if it did not, in some way, tap into something basic.

The second is to understand why normative replies to the skeptic seem so difficult to sustain. While we disagree with the skeptic's conclusion, the many attempts to reply to the skeptic all seem to fail. That failure, as Michael Williams has observed, is a failure to provide a therapeutic resolution to the problem - one that relieves us of our skeptical doubts by showing how the skeptic is mistaken in a way that is as clear and as intuitive as the skeptic's own argument.[121] Williams contrasts therapeutic diagnoses with theoretical ones, which aim to uncover the theoretical mistakes in the skeptic's underlying epistemological outlook. But, as Williams observes,

> If, as the New Sceptics[122] tend to think, scepticism presents an "intuitive" problem, a problem arising out of tensions in our most ordinary ideas about knowledge and justification, then burden of theory falls squarely on the critic, and theoretical responses fare badly against intuitive problems.[123]

Where Williams disagrees with the New Sceptics, and it's a disagreement that extends to the kind of position on offer by contextualists, is over what the apparent intuitiveness of the skeptical problem indicates. The kind of concessive response to skepticism on offer by contextualism is premised on the idea that because skeptical arguments are intuitive, this is indicative of the fact that they appeal to no more than our ordinary concept of knowledge and our ordinary epistemic practices. They further assume that because the skeptic's argument begins in ordinary practice, it does not rely on any deep or distinctive theoretical commitments. If, according to the contextualist, the skeptic does no more than take advantage of the context-sensitivity of "know," then the skeptic is not burdened by any theoretical presuppositions.

---

[121] M. Williams, 1996: xii - xxiii.

[122] Williams categorizes a number of contemporaries who see the problem of skepticism in Humean terms, as both intuitive and one we should be pessimistic about solving, as the "New Sceptics." He includes among them Barry Stroud, Thomas Nagel, P. F. Strawson, W. V. O. Quine, and Stanley Cavell.

[123] M. Williams, 1996: xvii.

This is what Williams denies. He argues forcefully, and correctly in my view, that intuitiveness does not indicate an "independence [from] problematic theoretical preconceptions." By way of analogy, he contrasts the paradoxes of relativity theory with Zeno's paradoxes of motion. While skeptical problems can "easily be made intelligible to people with no prior acquaintance with recondite philosophical theories" the paradoxes of relativity are not so easily detached from the large, counter-intuitive, and difficult theoretical framework of contemporary physics. But to show that paradox doesn't only rest on theory, Williams considers Zeno's paradoxes of motion, which can be explained quite easily, but also rest on deep and difficult questions about the fundamental nature of space and matter that are hotly contested.[124]

The contrast is instructive. That we do not consciously or explicitly undertake a commitment does not show that our judgments or intuitions about a subject matter do not reflect that tacit commitment. For instance, I have no idea whether space is infinitely divisible or not, and while I've considered the matter, I haven't considered it that hard. It does not follow from this that my judgments about the division of matter don't express a commitment to the infinite divisibility of matter. And this is because, plausibly, the many innate structures that subserve my tacit theory of physics, what some have called folk physics, encode this commitment in some way and this drives my intuitive judgments about divisibility. Whether or not the universe is organized that way doesn't matter, those systems evolved in an environment where the kinds of inconsistencies brought to light by Zeno's paradox never posed an adaptive problem. Likewise, our folk physics is not relativistic because the discrepancies between it and classical physics never emerged on a scale where the differences in their predictions made any adaptive difference to our ancestors. I'll explore these questions in much greater depth in the next chapter, but for here it is enough to outline the idea. Just as innate psychological structures shape our physical intuitions, so too should we expect that innate psychological structures shape our epistemic intuitions. Navigating the social environment was no less important for our ancestors than navigating the physical one, so we have every reason to think that our folk epistemology will exhibit the very same organization in advance of experience that our folk physics does. And if this is right, it's no less plausible to think that that folk epistemology will exhibit the same kind of

[124] M. Williams, 1996: xviii.

tacit commitments - commitments that may not be fully consistent. Understanding the geography of our epistemic psychology, I will note, is a very different project than the normative one, undertaken by Williams in *Unnatural Doubts*, of evaluating those commitments. In the first instance my goal is to describe.

But seeing this, we can see that perhaps contextualists and the New Sceptics have actually set themselves too difficult a task in replying to the skeptic. Because they've relied on intuitions about the skeptical problem gathered from the armchair, and they've failed to appreciate some of the variability in epistemic intuitions that has been uncovered by these studies, they have misdescribed the skeptical problem in a way that makes it more powerful than it actually is. Accepting that AI is intuitive may concede too much ground to the skeptic. But empirically describing the problem of skepticism can skirt the whole difficulty of weighing the force of contested intuitions. This should not be mistaken for the claim that because ordinary people do not find radical skeptical arguments intuitive we can reject the skeptic's argument. Williams' point operates in both directions: while some propositions we find deeply intuitive may rest on tacit commitments to incoherent epistemic commitments, some propositions we find deeply counterintuitive may rest on coherent epistemic commitments. The analogy with relativity theory, a consistent but counterintuitive position, illustrates the idea. The point is to detach the task of diagnosis from any kind of intuitive assessment of the merits of skepticism. One question is whether the skeptic's argument is a good one, another is whether we find it intuitive. But whatever the ultimate relation between our epistemic intuitions and their normative credentials, the descriptive project should be an empirical project, and by preceding the normative project, it sets the normative project on a firmer foundation. Taking up these questions, both normative and descriptive, will be the task of the next chapter, but before turning to those important inquiries, there are other implications of these results worth exploring.

One in particular is the role plausibility might play in resolving the debate between restrictionist experimental philosophy and more traditional approaches to philosophy that I canvassed in Chapter 1. Two of the most important results of the restrictionist program are that ordinary intuitions about epistemic thought experiments differ from those of trained philosophers, and that those intuitions seem to be highly variable in response to seemingly minor and irrelevant matters of presentation. Just to review, these include results that seem

to show that ordinary people do not share the philosopher's intuition about Gettier cases,[125] Fake barn cases,[126] True-temp cases,[127] Dretske's Zebra cases,[128] and even Brain-in-vat cases.[129] But in the philosophical literature these cases have been presented as abstract tests of principles, the primary function of which is to distinguish rival theories in an efficient way. Because good theories agree on so much, often times the cases used to differentiate them depart from the ordinary. The difference in responses between philosophers and ordinary people may largely be a matter of philosophers disregarding the implausibility of the scenario before issuing their verdict. Experiments 2a - 2d show that this is something ordinary speakers do not do. I think this is a very likely explanation of many of the discrepancies now discussed in the experimental philosophy literature.

Take, as an example, the Gettier scenario Starmans and Friedman experimented on: in it the focal agent of the vignette is taking a shower, and during this time a thief breaks into the agent's apartment, steals the agent's worthless plastic watch from a table, and then replaces it with an identical plastic watch, before leaving without stealing anything else; participants are then asked if the agent knows his watch is on the table.[130] The case is structurally the same as many Gettier cases, where bad luck is cancelled out by good luck, but Starmans and Friedman found that in this case ordinary people, unlike the philosopher, attribute knowledge. But this is subjectively speaking a very implausible scenario. The thief's behavior is difficult to rationally interpret, and so ordinary speakers might attribute knowledge to the agent in the vignette because they just ignore the implausible threat to that knowledge posed by an irrational thief with low ambitions. Compare these results to those reported by Jennifer Nagel, who uses a much more plausible stopped clock Gettier case. Nagel finds in these cases ordinary people do deny knowledge, and thus share the philosopher's intuition. Though most likely plausibility is not the only driver of differences

[125] Starmans and Friedman, 2012.

[126] See Calicoe, Machery, Buckwalter, and Stich, forthcoming.

[127] Weinberg and Alexander, 2006.

[128] Weinberg, Nichols, and Stich, 2001. It's worth noting that classically Dretske's Zebra is not a pure salience case because it is presented as a single case, where the evidence for the belief is immediately followed by the unrealized possibility of error, rather than as a pair of contrasting cases. I have used Dretske's Zebra case as a model for constructing contrasting salience cases here.

[129] Nichols, Weinberg, and Stich, 2004.

[130] Starmans and Friedman, 2012: 274.

in these studies, it is very plausible indeed that it is a large driver of difference, and when it is controlled for, perhaps much of the disagreement will disappear.[131]

Although now I'm speculating a bit more freely, it is also possible that the instability of epistemic intuitions, which is the focus of much of Jonathan Weinberg and Jason Alexander's work, may be a consequence of their using as their stimulus material the standard implausible thought experiments from the philosophical literature.[132] It may be that intuitions about cases become unstable when the scenario is strange, new, implausible, or the like, because it is in those cases that participants in experimental studies must contend with competing intuitions. Some, based in ordinary practice, might drive people towards one answer, while other principles, like indifference to implausible error possibilities, drive them in a different direction. In a case where strong competing intuitions are in play, it's reasonable to think that people's ultimate responses will be unstable. Perhaps people's epistemic intuitions will be much more stable when they are focused on plausible cases. If so, it will be one more reason why the early, and negative, programmatic conclusions of restrictionist experimental philosophy should give way to positive experimental philosophy.

Priming:

Experiments 3a & 3b showed that ordinary indifference to implausible error possibilities can be manipulated through priming. This priming effect can help illuminate what has been until now an ill-defined aspect of contextualist theories. The contextualists main hypothesis for explaining how contexts of epistemic evaluation change was via two phenomena, the introduction of an unrealized error possibility, and the rule of accommodation whereby standards shift in a way to make a speaker's assertion true. But contextualists tended to differ on how intuitive skeptical hypotheses are, and exactly under what conditions they are salient in a given context. They both needed to allow that the skeptic could alter contexts in a way to threaten knowledge, but avoid the result that just any wild yarn or conspiratorial whisper would do. In general, while clearly committed to the idea that the context of evaluation *can* change, contextualists have had little to say about *what causes* the change. Priming is one such cause. It can also help explain why we typically ignore the conspiracy theorist in ordinary life, but attend to the skeptic in the philosophy

[131] I am already conducting some of these studies, and the results from Fake barn cases are encouraging.
[132] See, for instance: Swain, Alexander, and Weinberg, 2008.

seminar room. One possible explanation is that, typically, conspiracy cases are not presented sequentially, first from plausible conspiracies to more and more implausible conspiracies. Usually the nuts go for the gold all at once. By contrast, radical skepticism in the form of AI is often presented in a carefully prepared way, both in the seminar room and in textbooks, because we are aware that many students initially have what DeRose has called the "Aw, come on" reaction.

Here the comparison between sequential presentation and mere repetition is important. Simply presenting cases of the same type over and over again does not seem have the power to generate salience effects with very implausible error possibilities. This is not to say that repetition *couldn't* achieve the same effect; it is certainly plausible that familiarity with a type of case could overcome our ordinary indifference to the implausible. Linguists, for example, sometimes develop a "dead ear" wherein they can no longer tell whether a specific sort construction is grammatical or not, simply because by repeatedly considering such cases their intuitive reactions are washed out. I think it's quite reasonable to expect that there is an epistemological equivalent of the linguist's dead ear, whereby careful and sustained study of cases, skeptical and otherwise, simply washes out any intuitive reactions we once had, including the heuristic filter I posit here, to explain our ordinary indifference to the implausible. In any case, for my purposes, what matters is the comparative claim. Cartesian priming through sequential presentation is *far* more powerful than mere repetition.

It is most unlikely that Cartesian priming is the only tool at the skeptic's disposal in preparing the ground for their case. And, I should emphasize, I in no way think that Descartes or the generic skeptic consciously deploys this mechanism. I believe it is a tacit form of a slippery slope argument that takes advantage of a general epistemic heuristic, to assess similar cases consistently. Nor do I think that contextual manipulations are the only important factor in diagnosing the appeal of the skeptic's position. As Figure 8 illustrates, the distribution of responses in the very implausible error case is bi-modal: some people deny knowledge in a fully-fledged way when confronted by the skeptic, others remain steadfast. This suggests that individual differences play an important role in the appeal of the skeptic's hypothesis. It is particularly in evidence when compared to the responses in the plausible error case, which are strongly right-skewed. Almost everyone feels the pull of plausible salience cases, but implausible ones are divisive. Although there were no

correlations between knowledge attribution and standard demographic measures like education, race, gender, or age, a future line of inquiry might explore whether the tendency towards skepticism is related to other stable, but more complex individual differences in cognitive style, including numeracy, need for cognition, need for cognitive closure, empathy, and big five factors like, extraversion, agreeableness, conscientiousness, emotional stability, and openness to experiences.

Another way in which these results illuminate the skeptical problem is that the make sense of it in personal terms. The typical way in which contextualists like Lewis, DeRose, and Cohen have approached the elusiveness of the skeptical threat is through conversational dynamics, and the metaphor of a conversational score. Invariantist error-theorist like Rysiew and Hazlett adopt a similar approach, but with a focus on pragmatics. In either case, while the starting point is a natural one for theorists focused on the linguistic features of our epistemic language, I think it encapsulates a basic misunderstanding of the skeptical threat. Surely we encounter skeptical doubts for the first time by hearing or reading a skeptical argument, but most often skeptical doubts are not *actually pressed on us by a skeptic.* The skeptical problem is not a problem being worked out in a conversation between two different individuals, each with different conversational aims and presuppositions, it is in the first instance a personal worry, one that we reflect on in isolation - as Descartes initially did. The dilemma is personal. The ambivalence is personal. And recognizing that the problem of skepticism is one we reflect on, the scope for the kinds of conversational misunderstanding that traditional contextualists appeal to is severely restricted. While we certainly can find opposing ideas attractive at the same time, we are just not so divided against ourselves. Tacit priming, on the other hand, especially of the kind that I posit here, can explain the personal significance of the skeptical problem. All that is required is that a person considers cases in a particular order, and they'll have altered the context of their reflection. If, indeed, the mechanism by which Cartesian priming works is through the consistency of assessment, that too is a principle that we deploy individually, and does not depend on contentious conversational dynamics. It is merely a consideration that has to be introduced, whether or not we are reflectively aware of it or not.

But this leads to a different question. Even if Cartesian priming can help alter the context of epistemic evaluation in such a way that the skeptic has can catch our ear, so to speak, why is it that the skeptic can't keep a hold of it? Once we accept the global skeptical

hypotheses, why don't we accept it for good? It's a good question, but not one that these experiments are designed to answer. One possibility is that the skeptic's argument in AI fails elsewhere, for example at the level of closure, as Fred Dretske and Robert Nozick have proposed. Perhaps while we accept that we don't know that a skeptical possibility of error doesn't hold, we nevertheless reject that knowledge is closed under known implication (i.e. we reject the idea that knowing that P, and knowing that P implies Q, it doesn't follow that we know that Q). According to this proposal, while the skeptic's first premise is true, you don't in fact know you're not a brain in a vat, it doesn't follow from this that you don't know that you have hands.

I don't think this is an option here, because standard salience cases like I've considered here rely on closure. So, participants are asked to indicate whether they agree or disagree that "John knows that the bird he saw is a Siberian Grebe". The unrealized possibility of error is that John hasn't considered that it's possible that he saw a Gadwell duck instead. Fitting this into the form of the argument from ignorance, we get:

1. John doesn't know he didn't see a Gadwell duck
2. If John doesn't know he didn't see a Gadwell duck, then John doesn't know that he saw a Siberian Grebe
3. Therefore, John doesn't know he saw a Siberian Grebe.

Because participants in the plausible error conditions tended to deny that John knew he saw a Siberian grebe, they in effect endorse closure tacitly. The denial of closure might play a role in explaining the bi-modality we observed in the Very Implausible Error case after Cartesian priming - perhaps closure is sensitive to plausibility - but it's unlikely to bear on the transience of the skeptic's conclusion.

So, the question of why skeptical doubt is so transient remains. Many have seen fit to echo Hume's observation that "Nature is too strong for principle" and that,

> The great subverter of *Pyrrhonism* or the excessive principles of scepticism is action, and employment, and the occupations of common life. These principles may flourish and triumph in the schools; where it is, indeed, difficult, if not impossible, to refute them. But as soon as they leave the shade, and by the presence of the real objects, which actuate our passions

> and sentiments, are put in opposition to the more powerful principles of our nature, they vanish like smoke…[133]

But while these researches don't speak to the transience of skeptical doubts, they needn't to make progress on the skeptical problem. The traditional diagnostic project has been to understand why skeptical arguments are intuitive. My reformulation of the task has been to ask under what conditions are different kinds of skeptical arguments intuitive to different people. But this is a different task than explaining why skeptical arguments fail to persuade us, even when they do strike us as threatening. Though related, how we reason about an argument and whether we update our beliefs based on that argument, are different matters. The transience of skeptical conclusions is a feature of how we update our beliefs, and understanding that is a different diagnostic project than the one I've engaged in here. A possibility I will take up in the final chapter of this project is that the failure to adopt the skeptic's conclusion, even when they've successfully manipulated the context in such a way that the argument strikes us as sound, is a feature of what psychologists call motivated reasoning (and everyone else wishful thinking).[134] The basic phenomenon is that there are certain kinds of beliefs, often self-assessments, that we systematically tend to favor when they cast us in a positive light, and there are certain kinds of beliefs we disfavor when the opposite is true. Almost everyone thinks they're above average in intelligence, driving ability, and conscientiousness. Indeed, 90 percent of college teachers believe they're above average at teaching. But none of this can be true. Half the people *have* to be below average. As I will argue latter, self-assessments of epistemic accuracy could certainly fall under this phenomenon. It is hard to believe that we know nothing of what we think we know; accepting the skeptic means accepting that we are deeply flawed epistemic agents, and perhaps that is just something we cannot do. Motivated reasoning - and here I'm speculating quite freely - may play an important role in explaining why Moore's argument seems so unpersuasive, even though it is so sensible. We feel a sense of tension when we accept an argument that is so self-serving, when we cannot find fault in the argument of our opponent.

But this is a thought I'll have to explore later. First I want to turn to *why* salience effects exist. If this chapter and the previous one have been about describing the skeptical

[133] D. Hume, 1978: Book VII, part II. See M. Williams (1996: 2 - 10), who calls this bi-perspectivalism.
[134] Z. Kunda: 1990.

problem, the next will be about offering a fuller diagnosis of it based on that description. In the next chapter I will argue that salience effects are a product of competing epistemic systems. I'll defend this view, which I call psychological contextualism, and then I'll turn to the question of whether or not the diagnosis of the skeptic's argument on offer here can be the basis of an assessment of that argument. Should we endorse the skeptic? If her argumentative strategy were fully explicit, would we endorse it?

Chapter 5

The Psychological Foundations of the Skeptical Problem

a man always believes more readily that which he prefers

- Francis Bacon

1. Introduction

Can our intuitions about how to use the word "know" shed any light on the problem of Cartesian skepticism?

For the last half century or more, the consensus of an important philosophical tradition has been to answer "yes." In ordinary language philosophy, and indeed in analytic philosophy more generally, linguistic analysis has often aimed at dissolving philosophical problems by showing that they rest on linguistic confusions. The prevailing approach to understanding the skeptical problem over this period is an important example of this tradition in action. Linguistic contextualism as we've examined it here[1] is the most prominent and important contemporary example of a tradition of inquiry that stretches back through J. L. Austin, Norman Malcolm, Ludwig Wittgenstein, and even to G. E. Moore and the birth of analytic philosophy.

I believe this tradition has delivered many insights, but that it has also obscured some possibilities, perhaps in part by too quickly assuming that our language is also the language of thought. The alternative that I want to sketch here is that the skeptical problem has a deeper, psychological foundation. My proposal is that the skeptical problem is a byproduct of different psychological systems applying inconsistent standards of evaluation to related claims. If I am right, the conflicting intuitions at the heart of the skeptical problem are a byproduct of our evolved epistemic psychology, and not the result of linguistic confusion. Linguistic intuitions about "know" are a vital kind of evidence, but not because the skeptical problem is a byproduct of linguistic phenomena. Rather, I will argue that our intuitions about "know" are an important source of evidence about our epistemic psychology. Any adequate diagnosis must look at the lower, more fundamental level, if it is to have any hope of explaining the intransigence of the skeptical problem.

In the following I'll first consider the philosophical significance of the skeptical problem, and give an account of the dialectic confrontation between the skeptic and anti-skeptic. I'll use this to develop a set of commonsense desiderata for an adequate diagnosis of the skeptical problem. I will then develop an account of the components of our epistemic psychology that give rise to the skeptical problem, paying special attention to how the conflict of intuitions at the heart of the problem are tokens of much more common

[1] Advocates of which, as we've seen, include, N. Malcolm (1952); G. Stine (1976); D. Lewis (1996); S. Cohen (2001); and K. DeRose (1995).

types of epistemic judgments. I will then compare linguistic contextualism with the psychological theory I favor, evaluating both against the desiderata developed in the preceding section and the empirical results reported in Chapters 3 and 4. I will argue that the psychological theory is far more successful in accounting for the data, and has the further virtue of offering a more substantial account of the skeptical problem while nonetheless remaining *consistent* with linguistic data that contextualists advert to in their explanations. A key result is that skeptical arguments like AI are only successful because they appeal to a cognitive bias for consistency. I will conclude by arguing that the empirical description of the skeptical problem undertaken here does provide the resources for offering a reply to the skeptic that relieves us of the pressure to reply to skeptical arguments like AI. For, if it is true that the cognitive significance of the skeptic's implausible error scenarios tacitly depends on priming a cognitive bias towards consistency, then it becomes clear that the skeptic's argument is not as presumptionless as it at first seems. As many have recognized, the skeptic must be careful to present their arguments in such a way that they do not become self-refuting, and it is in this way that AI has been such an important formula for the skeptic: the philosophical consensus accepted it as *intuitive*. But, if there is a step in the skeptic's argument that must be defended, then the philosophical force of radical skepticism is muted.

In many ways the ultimate conclusion of this project is about the power of empirical investigation in responding to the skeptic. It is empirical observation that reveals that the AI depends on a tacit bias. And so, we can see the power of genuine empiricism: by looking to the world to answer our questions we can even answer the most powerful challenge levied against that method.

2. The Skeptical Problem

Why, ultimately, is the skeptical problem a problem? It seems to reveal a deep confusion in our ordinary notion of knowledge, and in our ordinary thought and talk about knowledge. On the one hand we speak of knowing many things - what day Christmas falls on, when the car will be fixed, and where we were born - and this way of speaking is automatic and intuitive. On the other hand, skeptical arguments seem to present an intuitive case that we know very little, if anything, at all. We feel the tension of simultaneously wanting to affirm that we have knowledge of ordinary facts, and at the same time feeling

drawn to the conclusion of an argument that denies any knowledge of ordinary facts. But then we're left with the question: do we know a lot, or nothing?

Up to this point I have argued that AI is not as intuitive as the received view has been willing to accept. I have tried to empirically demonstrate that the appeal of skeptical arguments is much more limited than previously thought: *some* people, in *some* contexts, find *some* skeptical hypotheses threatening. This is not to foist onto the opposition the implausible idea that we are all paralyzed by doubt once we encounter an argument like AI. This is far from the case. Many have been happy to point to Hume when considering the transitory force of skeptical doubts.

> Nature is always too strong for principle…the first and most trivial event in life will put to flight all doubts and scruples, and leave [us] the same…[as] those who never concerned themselves in any philosophical researches.[2]

But Hume's point concerns how doubt is conditional on attention, and that even small practical concerns can draw us away from philosophical reflection; he does not seem to mean that practical concerns are a way to allay those doubts permanently. What is novel about the observations I've made is that they show that there are robust individual differences in the threat we perceive skeptical challenges to pose to ordinary knowledge, even when it has our full attention. The problem, then, is not merely that skeptical arguments like AI threaten all empirical knowledge, or that all people find skeptical arguments threatening.

Neither is the problem that the skeptic raises standards simply by asking, "are you absolutely certain?" Nor even by issuing a challenge "can you rule out all possibility of error?" High-standards skepticism, while venerable, does not accord intuitively with our ordinary ways of talking and thinking.[3] Moreover, high standards must be defended, and once the skeptic is in the game of offering justifications, their own argument works against them.

Some believe that what makes Cartesian skepticism truly radical is that it aims to undermine the possibility of epistemic assessment by undermining our ability to distinguish between beliefs and anything better. This worry is more theoretical, and it is deeply

[2] D. Hume, Enquiries, Section XII, part 2.

[3] See Alan Bloom's translation of Plato's *Republic* for a high-standards view: "could any man count that which doesn't make mistakes the same as that which does?" (477e). See S. Nichols & A. Pinillos (forthcoming) for a related discussion.

connected to the normative project of giving a positive account of knowledge. It is not my worry, however, because my project is not to give a normative account of knowledge.

My concern, and to my mind what makes arguments like AI threatening, is that they are a simple formula for undermining most beliefs, irrespective of the evidence we have for them, how reliable they have proved to be in the past, and regardless of how the world is or how those beliefs cohere with everything else we know. Take an empirical claim, no matter how familiar or recently grounded, think of an unrealized way in which you might have that belief falsely, and you have the materials to undermine our knowledge with the AI method. That our knowledge is generally vulnerable to a formulaic procedure *virtually unconstrained by how the world actually is,* is what makes skeptical doubts so problematic: why do we sometimes find them so compelling despite their lack of connection to the world?

Indeed, it is this disconnection from the world that helps underwrite their normative force. Cartesian arguments like AI are (according to the received view) intuitive: they are presumptionless, and because they seem to appeal to nothing more than our ordinary ways of thinking and talking, they need no defense. Michael Williams aptly describes the tactic. The skeptic seems to use an argument that "does not depend on contentious or elaborate theoretical ideas about knowledge and justification."[4] In this way the skeptic occupies the high ground, and places the burden of proof onto the anti-skeptic.

Because my project is descriptive and empirical, part of the challenge of making sense of skeptical doubt is to explain how our evolved psychology would allow for such generic conditions of defeat to any belief. It is a standard assumption of evolutionary biology and evolutionary psychology that the greater the importance of a feature of the phenotype to the survival of the individual, the stronger the selective pressure will be upon it. Over the course of our evolutionary history, acquiring knowledge of the world around us was certainly among the most important tasks in navigating it, and so we can assume that there was a strong selective pressure on the mechanisms devoted to that task. By no means is every feature of epistemic assessment the product of processes under selective pressure.

[4] M. Williams, 2001: 5. Williams goes on to offer one of the most revealing accounts of the skeptical dialectic "If, as the New Sceptics think, skepticism presents an "intuitive" problem, a problem arising out of tensions in our most ordinary ideas about knowledge and justification, the burden of theory falls squarely on the critic, and theoretical responses fare badly against intuitive problems…to make headway against the sceptic, then, the critic needs first to level the playing field. It may not be necessary to shift the burden of theory entirely to the sceptic's shoulders. [sic] but it is crucial that the sceptic be forced to acknowledge his share" (1996: xvii -xviii).

However, the more systematic the feature, especially when it is not the object of explicit instruction, the more likely it is to be the end result of selective pressure. Salience effects are one such example: they are systematic, but not the product of explicit instruction. Even though evolution is a satisficer and not an optimizer, we should aim for an account of the skeptical problem that does not, in the end, force us to conclude that it is an ecological irrationality, an unmarked and unexplained cul-de-sac in our mental geography. In short, the problem of the skeptical problem is to make it make sense. To explain how it arises and why it arises, and to do so, if possible, without stopping at some level and submitting that this is the level where a mistake happens.

The specifics are by now familiar. As I have described the problem of the preceding four chapters, AI rests on three intuitions, which are naturally compelling individually, but jointly incompatible. We have the *Ordinary intuition*: that most ordinary knowledge claims are true. We also have the *Cartesian intuition*: that skeptical hypotheses undermine our ordinary knowledge. And finally, we have the *Intuition of conflict*: that the conclusions of skeptical arguments contradict ordinary knowledge claims. Diagnosing the problem doesn't mean developing a response that would convince a skeptic to withdraw their challenge. Rather, its aim should be to shed light on the appeal of these intuitions. So much the better if it can, to use Jim Pryor's phrase, "puts our minds at ease."[5] As we saw in Chapters 1 and 2, this formulation of the problem, in terms of the argument from ignorance, and this approach, in term of intuitive diagnosis, are both widely held.[6]

A related element of the skeptical problem that deserves special attention is that it is personal. The skeptical problem is one that generates first-person *ambivalence* in the classical sense of the term; it generates a feeling of strong, simultaneous, attraction to the intuitions that *we do have knowledge of the external world* and that *we do not have any knowledge of the external world*. The intuitiveness of these statements is personal in the sense that one does not meet an actual skeptic and struggle to refute the argument they put to us; rather, we all, on occasion, struggle to come to terms with the skeptic within ourselves. It is not an argument between two people, but a personal confrontation.

This outline of the skeptical problem dictates the agenda for responding to it. To reiterate, the philosophical significance of skeptical arguments like AI is not that they result

[5] J. Pryor, 2000: 517. I will argue that putting our mind "at ease" is likely impossible.
[6] S. Cohen, 1999: 83; D. Pritchard, 2002: 217; P. Rysiew, 2001: 480.

in global skepticism about the external world, but that they offer a formula for undermining any particular piece of perceptual knowledge without having to track how the world actually is in fact. What we should want to understand is why our psychology is such that this is possible. More than putting our minds at ease, an adequate account of the skeptical problem should explain how such an inconsistency in our thinking can arise, how it can be so pervasive, and how it fits into the larger story of our coming to know about the world. What we should want from an account of the problem is an explanation of how this is possible. The empirical approach I pursue here can succeed where past attempts that have aimed, as I've said, "to put our minds at ease," have failed. Such attempts have failed in the past because they begin from the assumption that the problem is the inescapable product of our concept of knowledge. But as we will see, some habits of thought, like cognitive biases, are intuitively attractive but normatively fallacious. So we can distinguish diagnosing the source of the intuitions behind arguments like AI, and offering an evaluative diagnosis of the skeptical problem. Empirical investigation aims at the former, but it can then be used as a basis for the latter. Empirical investigation can put our minds at ease not by trying to remove a troublesome intuition, but by positioning us to evaluate the principles that give rise to them. By empirical observation we can move past simply trading intuitions, or speaking metaphorically about their force. In what follows I set out some desirable goals for an account of this kind of skeptical problem.

3. Desiderata of a Diagnosis of the Skeptical Problem

In the broadest terms, the puzzle posed by the skeptical problem is to explain how we find two inconsistent propositions attractive. A good diagnosis will offer some understanding or explanation of this attraction. Explanation can be a vague requirement but I propose the following as desiderata, in order of importance, of any diagnosis of the skeptical problem.

i. It should be empirically adequate

> Any genuine diagnosis must explain the first person ambiguity the skeptical problem provokes. It must explain, from the perspective of a single individual, why our knowledge of the external world seems so undeniable, and why the conclusion of the argument from ignorance seems so irresistible.

This will involve explaining the source of the core intuitions behind the skeptical problem. First, it should explain the intuitions behind salience cases. In particular the theory should explain why introducing an unrealized possibility of error seems to intuitively undermine a person's willingness to attribute knowledge. It should explain why this willingness is inversely proportional to an individual's judgments about the plausibility of the error. Furthermore, it should explain why sequential presentation overcomes some individual's indifference to the implausible. Finally, it should explain the intuition of conflict. It should explain why some individuals experience it, while others don't.

ii. It should be theoretically unified

A successful diagnosis will be theoretically unified. In offering an account of how different factors affect individuals' responses to the intuitions behind AI, explanations in terms of established and general phenomena are (ceteris paribus) preferable to those that must posit novel or specific phenomena.

iii. It should be charitable

Theories that make sense of individuals are preferable to theories that attribute error. The platitude that supports this claim is just that uncharitable theories are too easy to overgeneralize.

A successful diagnosis will avoid positing *sui generis* errors to individuals. We want an explanation of the problem that actually *explains* the rational conflict, and does not merely *displace it*. Charity involves two sub-claims, that positing error is less preferable to not, and positing specific types of error are less preferable than general types.

Theories that attribute errors of a general *type* are preferable to theories that attribute errors of a local *type*. The terms "general" and "local" deserve some elucidation. By "general errors" I don't mean theories of global error: this would conflict with the immediately preceding claim that we should prefer theories that vindicate our background beliefs. Instead by "general" I mean errors of a recognizable *type* are preferable to errors that occur with only a single type of term, or in a single type of case.

Thus, a diagnosis that can explain the skeptical problem without attributing systematic error is preferable to one that does; and, a diagnosis that can explain the skeptical problem as a general kind of error is preferable to a theory that must posit a local type of error that only occurs in the skeptical case.

Perhaps it is also worth mention one reprimendus:

i. It need not solve the skeptical problem

> If I were to suggest a diagnosis of why many linguistic diagnoses have failed to offer an adequate analysis of the skeptical problem, it would be that they have all been driven, deep down, by the goal of refuting or somehow resisting the skeptic's conclusions. And as a consequence they must at some point or in some place attribute error to speakers. Ultimately, however, I think it has tempted them to attribute rather implausible kinds of error.
>
> A genuine diagnosis of the skeptical problem - of why we find skeptical arguments so convincing - need not provide the means to solve the problem. Perhaps the skeptic is right, in which case preserving our common knowledge is no virtue.
>
> Here an analogy is perhaps useful. Zeno's paradoxes of motion raise a metaphysical question, is space infinitely divisible or not? Now, it seems like, as a matter of metaphysical necessity, that space is or isn't, but not both. But it is not a virtue of an explanation of *why* we find Zeno's paradox compelling that it favors one or the other of the solutions. It just needs to explain why we find both conceptions of divisibility so attractive. How things are is another question entirely.
>
> So, diagnosing the skeptical problem need not tell, one way or the other, in favor of a particular approach to knowledge.

An obvious word of caution in applying these desiderata is in order. Explanation is not wholly objective. This is true in the empirical sciences and it is equally true, too, in this epistemological inquiry. The conditions laid down here are also, obviously, not objective. What they aim to do is clarify the general principles of responsible inquiry in a field, nothing more and nothing less. The truth provides the best explanation, regardless of whether we happen to know the truth or not. And here what we are looking for is just what it is that causes beings like us to be feel attracted to two incompatible propositions.

What *feels* like the best explanation, however, isn't necessarily the truth. What we find to be the most convincing explanation is often the one that leaves our background beliefs undisturbed. But sometimes this is a mistake because we have false background beliefs and we must reject what is seemingly intuitive. The word of caution, then, is that these desiderata are features conducive to truth, but not constraints on it. This is an apt opportunity, then, to raise a useful distinction introduced by Michael Williams: there is a

meaningful difference between a therapeutic response to the skeptical problem and a theoretical one.[7] A therapeutic diagnosis attempts to relieve us of our skeptical doubts once and for all: it aims to reveal a flaw in the skeptic's position in such a way that we are no longer troubled by radical doubt.[8] But this might not always be possible. Returning to example of Zeno's paradox. It may be that our means of spatial cognition, specifically the many systems that subserve it, and our concept of space are not consistent. Moreover, there is little reason to think they reflect the actual physical reality. If this is the case, it may never be possible to relieve the intuitive conflict Zeno's paradox engenders. It may only be possible to develop a theory of space (or division) that is useful and accurate, even if it is counterintuitive. So, we can distinguish three projects. One is diagnosing the source of the intuitions that give rise to the skeptical problem. Another is giving a solution to the skeptical problem that shows we needn't affirm one of those intuitions. A third is trying to remove the pull of the intuitions.

But this, also, helps illuminate the power of the empirical approach I have adopted here in trying first to describe, and then diagnose, the skeptical problem. To put it simply, I believe a heap of empirical evidence can succeed where therapeutic diagnoses fail. A heap of empirical evidence can relieve of us the need to try to accomplish the third project above. The probative force of therapeutic diagnoses depends on personal appraisals of the strength of competing intuitions. The probative force of an empirical diagnosis does not. It doesn't matter whether the evidence is counterintuitive, just that it is accurate.

4. Psychological Contextualism

The proposal I wish to sketch is that the skeptical problem is a byproduct of conflict between the outputs of different processes in our epistemic psychology. The overarching result of the last three decades of research into cognitive psychology is that there are

[7] M. Williams, 1996: introduction.

[8] Another example is the Monty Hall problem. A familiar game show challenge involves presenting a participant with three doors, one with a car behind it and the other two with goats behind them. Once the participant chooses which door to open, the host opens one of the two remaining doors, always revealing one of the goats. At this point the host asks the participant whether they would like to switch or not. The untutored intuition is that it doesn't matter probabilistically if the participant switches, because it's a 50/50 chance either way. This, however, is wrong. The participant should always switch, as the unselected door has a 2/3 chance of hiding the car. This seems wrong, initially, but becomes much clearer when it's explicitly described using a frequentist interpretation of probability. Compare your intuitions in the standard Monty Hall to those in Big Monty Hall: the set up is as before, but now there are 100 doors, 99 of which hide goats. After making your initial selection the Host opens 98 doors, all hiding goats. Now, is it better to switch, or stand pat?

multiple, domain specific processes for reasoning that employ a variety of specific, and sometimes inconsistent standards of evaluation.[9] Sometimes known as the heuristics and biases tradition, the view holds that the function of reasoning is to produce ecologically rational beliefs about the world, where an ecological rationality belief is understood as one that leads to adaptive behavior.[10] The view is often contrasted with an older tradition in philosophy and psychology, sometimes called the Cartesian view, according to which reasoning depends on a single, conscious, deliberative process that aims at true beliefs, and that when functioning properly involves a single set of consistent standards of evaluation.[11] While this later view certainly has a role in normative theorizing, it no longer is a serious competitor to the accumulating empirical evidence about cognition.

According to the view I favor, we have two psychological systems, or perhaps clusters of systems, for evaluating potential beliefs. These systems have different functions, and thus apply different standards of evaluation, based on certain individual and contextual factors. One, that I will call the permissive system, employs a standard of evaluation biased toward confirmation, and functions to confirm or validate potential claims by searching for evidence that is consistent with them. A second, that I will call the stringent system, employs a standard of evaluation biased toward disconfirmation, and functions to dismiss potential claims either by searching for evidence that is inconsistent with them, or by searching for circumstances otherwise incompatible with them.[12] In essence, the first system asks *can I believe P?*, while the second asks *must I believe P?*[13] On the view I will defend, these two systems underwrite the conflict at the heart of the skeptical problem. A neutral claim

[9] See, for example, E. Stein (1996), G. Gigerenzer (2002), and D. Kahneman (2011).

[10] The name "heuristics and biases" tradition may sometimes conjure up a picture of a research program focused on identifying local and idiosyncratic deviations from a neutral, domain general system that is itself very similar to Cartesian theory of reasoning. While those involved in the heuristics and biases may once have thought that that was what they were doing, it is no longer accurate. As we will see, there is a very large body of empirical results that suggests that even when reasoning takes the form of conscious deliberation, it is nevertheless subserved by several domain-specific strategies. It is not that reasoning sometimes exhibits biases, but that reasoning is biased information processing.

[11] D. Sperber, 2001; D. Sperber and H. Mercier, 2011.

[12] I do not mean to posit that there are only two systems, or that these two systems mark the major dichotomy in of epistemic psychology. Rather, I only mean to claim that they represent two very important strategies.

[13] Tom Gilovich frames the research on motivated cognition in terms of these two questions (1993: 84). See also Dawson, Gilovich, & Regan (2001: 1380). As we'll see, the literature on motivated cognition plays an important role in the theory I sketch here. However, Gilovich proposes these questions as specific heuristics that we follow when reasoning about a subject about which we have a preferred conclusion. I think the questions express a general set of strategies of evaluation, not mere heuristics, that are sensitive not only to the content of a conclusion, but to the form in which it is presented, and the source producing it.

like "I have hands" is normally evaluated against the *can I* standard, and leads to the ordinary intuition. However, skeptical arguments prime us to deploy the *must I* standard, and this leads to the Cartesian intuition.

The theory on offer differs from linguistic contextualism in a number of ways. Linguistic contextualism, like we see in the work of Keith DeRose and Stewart Cohen uses scalar adjectives as a model for the semantics of "know," and an assumption of their view is that epistemic standards change by degrees along a continuous scale, much as measures of flatness can come by degrees. By contrast, the view I want to defend here is that contextual factors and individual differences prime us to use categorically different standards in evaluating potential claims. Moreover, where linguistic contextualism is a view tied to the behavior of a single term in virtue of its semantics, the view on offer here is general, and applies to a host of epistemic attributions, including knowledge, belief, justification, certainty, probability estimates, estimations of understanding, and degrees of explanation.[14] Where epistemic contextualism makes a local prediction about the behavior of a single term, the psychological account I am developing here predicts that categorically different patterns of evaluation will emerge across our entire epistemic repertoire.

So far, however, I have only offered a sketch of the diagnosis of the skeptical problem. I have not offered any evidence for the proposal, or any specifics about how skeptical arguments like AI prime us to deploy stringent standards, or what those standards are. In what follows I will summarize a large body of empirical work in support of the view. First I will summarize the literature on confirmation bias, and explain how it involves the application of a broad, permissive standard of evaluation that has the general form of a fallibilist theory of knowledge. Then I will review the literature on disconfirmation bias, and explain how it shows that across a wide variety of epistemic evaluations we deploy more stringent standards that have the general form of infallibilist theories of knowledge. I will then review the phenomenon of motivated cognition, and explain how the content and source of a claim affects the context-sensitive engagement of these systems, and the conditions under which they are activated. In the course of this I'll explore the argumentative theory of reasoning, and offer some speculative remarks about how our

[14] Though I will not reproduce the experiments here, in a forthcoming paper with Josh Alexander and Chad Gonnerman, I show that it is possible to generate salience effects for "belief," "justification," "explanation," "understanding," and "likely." Salience effects are everywhere.

evolutionary trajectory led to both the permissive and stringent systems of evaluation. After this I'll give a detailed account of how these systems lead to the competing intuitions at the center of the skeptical problem, and how they fit with the empirical results from Chapters 3 and 4. I'll conclude the section by considering two possible challenges to the account.

4.1 Confirmation Bias

Confirmation bias is a widely studied feature of cognition whereby information is processed in such a way that it becomes unlikely that a focal hypothesis, claim, or argument will be disconfirmed or otherwise dismissed. The process has been studied under many names, including congruence bias,[15] my-side bias,[16] and positive testing strategy.[17] A broad array of studies across a number of domains have shown that the ways we search for information, the ways we remember information, and the ways we test claims, all function in a way to "foster the immunity of [a] hypothesis."[18]

The process is possibly the most familiar cognitive bias, and it is easy to illustrate. Instead of forming two or more plausible hypotheses to explain a state of affairs, we tend instead to search for evidence that is consistent with a standing account (what I call a focal hypothesis). In one experiment, participants were informed that one of two buttons operated a light, and they were asked to formulate a hypothesis about which button it was. When they were subsequently asked to test their hypothesis, overwhelmingly, participants chose to test their hypothesis by pressing the button that they had guessed operated the light. This is despite the fact that pressing either button, i.e. testing the focal hypothesis or testing the counter hypothesis, would effectively determine the truth. In the absence of the bias, we might expect that participants would be no more likely to prefer one method of testing to the other.[19]

A number of experimental paradigms reveal confirmation bias. Klayman and Ha have found evidence for it as a general learning procedure in studies of object categorization.[20] Richard Nisbett and David Ross have found evidence of it in evaluations

[15] J. Baron, 2008.
[16] J. Baron, 1995; M. Toplak, and K. Stanovich, 2003; M. Toplak, R. West, and K. Stanovich, 2013.
[17] P. C. Wason, 1960.
[18] M. Oswald & S. Grosjean, 2004: 79.
[19] J. Baron, 2008.
[20] Klayman & Ha, 1989.

of the strength of correlations.[21] But the first, and perhaps most famous example, is the Wason selection task. It turns out, as we will see, to be one of the more instructive as well, illustrating two important features of cognitive biases. In the classic case, individuals are presented with the task of testing a rule like the following: if a card has a vowel on one side, then it has an odd number on the other side. They are then presented with an array of four cards, and told that each card has a letter on one side and a number on the other. They are then told to flip over all and only the cards that are necessary to test the accuracy of the rule.[22] To illustrate, suppose the array included cards displaying a 7, an N, a 2, and an E. The task has been exhaustively studied in a number of domains, but typically, when the task is presented as I have presented it here, more than 80% of individuals would turn over the E card because it shows a vowel. And while this choice could disconfirm the rule (if it showed an even number) it is also a way to find evidence for the rule. Likewise, the next most common choice is the 7 card, because it also could serve to provide evidence for the rule, despite the fact that it *cannot* provide evidence to disconfirm the rule. The choice of cards that can actually provide evidence against the rule are the E card and the 2 card, but typically fewer than 5% of individuals chose this combination.[23]

What is perhaps most instructive is that direct instruction does little if anything to alter how people approach the task. A number of follow-up studies have attempted to find effective debiasing measures, but none seem to have a robust effect.[24] Indeed, the task really only asks people to evaluate a statement expressing a material conditional, and yet a study that compared performance on the task before and after participants took a semester long course in formal logic found that direct instruction only reduced the rate at which individuals made at least one error by 3%.[25]

I do not want to make the task sound impossible! It definitely is not. With a bit of effort and reflection it is more than possible to learn how to respond to it correctly. Nevertheless, these studies have two important lessons over and above what they show about confirmation bias and direct hypothesis testing. First, they show that confirmation bias is very general: the Wason selection task has been studied across a number of subject

[21] Nisbett & Ross, 1980.
[22] P. C. Wason, 1968.
[23] P. Johnson-Laird & P.C. Wason, 1970.
[24] P. C. Wason, 1969; Wason and Shapiro, 1971.
[25] Cheng, Holyoak, Nisbet and Oliver, 1986.

matter domains, and with very few exceptions the same pattern of biased evaluation shows up across the board.[26] Second, the bias is tacit and intransigent: as we've seen, the intransigence of the task performance in the face of direct instruction demonstrates that the processes leading to the typical performance are tacit and deeply ingrained. Confirmation bias in the form of the Wason selection task exhibits features associated with cognitive biases more generally, i) they are involuntary and do not depend on conscious instruction or willful direction, and ii) they are hard though not impossible to avoid.[27]

4.2 Confirmation Bias, Disconfirmation Bias, and Motivated Reasoning

Disconfirmation bias is the counterpart to confirmation bias. It is the process or set of processes whereby information is processed, retrieved from memory, or otherwise evaluated in such a way such that a focal hypothesis or argument is likely to be disconfirmed or dismissed as unsupported. Confirmation bias and disconfirmation bias are examples of more general phenomena, known as motivated reasoning, and cannot be properly understood independent of it. Motivated reasoning, sometimes also called directional bias, is an overarching, tacit information processing stance whereby an individual's attitude toward a focal hypothesis or argument determines the procedures by which it is evaluated or how memory is searched for relevant information.[28] Hypotheses and arguments that are abstract, or toward which an individual has a neutral or positive attitude, or that come from a trusted source automatically activate what I have called the permissive system, and thus give rise to confirmation bias. Hypotheses and arguments toward which an individual has a negative attitude, or which come from an untrusted source, automatically activate what I have called the stringent system, and thus give rise to disconfirmation bias. By prior attitude I mean that an individual has a prior conclusion about the subject of a newly introduced claim or argument, or that an individual will be portrayed in a negative or positive way by a newly introduced claim or argument. P. H. Ditto and D. F. Lopez sum up this tendency to apply different standards of evaluation to information depending on its source or content in the

[26] Mynatt, Doherty, & Tweney, 1977; L. Cosmides, 1989; M. Oswald & S. Grosjean, 2004.

[27] Pohl & Hell, 1996: 3. These features are also often associated with optical illusions, although they are much weaker for cognitive ones. For some illusions, a proper instruction, careful selection of the material, or other procedural variation may reduce or even eliminate the illusion (as an example, see Gigerenzer, Hertwig, Hoffrage, and Sedlmeier in press), while for other illusions, most (if not all) attempts to overcome the effect have failed (as an example, see Pohl & Hell, 1996).

[28] For a wide-ranging review of motivated reasoning, see Z. Kunda, 1990.

following way, "people are less skeptical consumers of desirable information than undesirable information."[29] It deserves emphasis that these biases are the product of processes that apply across a number of domains, including i) hypothesis testing and formation, ii) memory search, iii) explanation evaluation, and iv) are sensitive to both the source and content of information.

Confirmation bias and disconfirmation bias are two sides of motivated reasoning, not two independent phenomena. Confirmation bias was likely discovered and described first because the first studies of hypothesis testing focused on abstract problems, like the classic Wason selection task, or focused on neutral subjects like the light bulb test, that did not challenge or support any of the participant's prior convictions. I mentioned before that there are relatively few interventions that affect performance on the Wason selection task; one that can alter performance is to use conditional statements about which the experimental subjects have strong prior beliefs. It turns out that when the rule to be tested supports a strongly held prior belief or supports a conclusion they find agreeable, participants tend to use the positive test strategy discussed above. However, when the conditional statement challenges a strongly held belief, or supports a conclusion they find disagreeable, they are much more likely to use a negative test strategy, and choose the cards that would, in fact, disconfirm the rule. Erica Dawson, Thomas Gilovich, and Dennis Regan used the Wason selection task to compare how individuals reasoned about death in the abstract and how they reasoned about their own death. In a study with two independent conditions, participants reasoning about death in the abstract performed at the normal level on the task, correctly selecting the cards that would test the rule between 5% and 10% of trials. However, participants who were confronted with a rule they find highly disagreeable, about the prospect of their untimely death, were much more successful on the task, selecting the appropriate cards approximately 45% of the time.[30] Their work shows that when a claim or argument has neutral implications for an individual, they tend to search for information that confirms it. But, when the rule has a conclusion they find aversive, they more often are

[29] P. H. Ditto & D. F. Lopez, 1992: 568.
[30] E. Dawson, T. Gilovich, and D. Regan, 2002.

able to generate a counter hypothesis incompatible with the truth of the rule, and search for evidence of the truth of the counter hypothesis.[31]

Selective memory search also leads to motivated reasoning. In the fictional town of Lake Woebegone, from Garrison Keillor's PRI radio program *Prairie Home Companion*, "all the men are strong, all the women are good looking, and all the children are above average." The lake woebegone effect is one of the familiar examples of motivated reasoning, and takes its name from this tagline. When people are asked to evaluate themselves, they seem to consistently overrate their own performance. For example, in a study of college professor's self-assessment of teaching ability, 94% rated themselves above average.[32] The same kind of positive bias in self-assessment has been found in studies of i) intelligence, ii) fairness, iii) prejudice, iv) driving skill, v) leadership, and vi) future prospects. The phenomenon was originally studied in the context of self-assessments of individual contributions to group work, and was initially called self-serving bias.[33] The leading contemporary account is that motivated reasoning in the form of biased memory search leads to these results. We evaluate a claim that presents us in a positive light be searching our memory for evidence that confirms it (I have had no speeding tickets in the last 12 years) rather than evidence that would challenge it (I have been in three accidents in the last 12 years).[34]

The same search bias appears toward claims about which we are neutral, as illustrated by Snyder and Cantor's study of personality assessment. In their study participants were asked to read a biography of an individual that was seeded with facts that were evidence for extroversion and an equal number of facts that were evidence for introversion. Several days latter the participants were recalled and split into two groups: one group was asked to judge whether the hypothetical candidate was suitable for a job requiring extroversion, the other was asked if the person was suitable for a job requiring introversion. In both cases, participants were much more likely to report facts showing the person was

[31] As I'll argue, this is highly significant. It is only in certain circumstances that we generate and evaluate claims by comparing them to counter hypotheses. I will argue that salience effects, at the heart of the contextualist literature on skepticism, are an example of this process.
[32] P. Cross, 1977.
[33] D. T. Miller & M. Ross, 1975.
[34] Z. Kunda, 1990; T. Gilovich, 1993: Chapter 5.

suitable for the position they were asked about. This result supports a biased search strategy.[35]

Claims that portray us in a negative light, or have aversive conclusions can trigger a selective evaluation in the opposite direction. In a classic experiment on evaluations of the effects of tobacco use, Kassarjian and Cohen showed that nonsmokers find studies supporting the health risks of smoking more persuasive than do nonsmokers.[36] This leaves open the possibility that prior beliefs about tobacco use influence updating, rather than pure motivation, and so it does not demonstrate, on its own, that individuals are applying different standards in evaluating a particular claim or body of evidence. But, in a study of the evaluation of caffeine use, Ziva Kunda found that women who are heavy consumers of caffeine are more likely to reject the results of a purported study showing that caffeine is risky for women than are women who do not consume very much caffeine. The results parallel the previous study, but interestingly, she found that men who are heavy consumers of caffeine, and thus likely to have the same prior beliefs about the health effects of caffeine consumption, are not more likely to reject the study.[37] Because this study attempts to control for the effect of prior beliefs, it is important evidence for the conclusion that individuals apply different standards in evaluating a particular claim.

The earliest accounts of motivated reasoning suggested that results like these were examples of self-deception, or defensive inattention to relevant facts. But recent results do not show that individuals are aware of being biased. An alternate theory via Arie Kruglanski is that motivated cognition is a product of depth of search. Because epistemic inquiry has no fixed stopping point, and because our cognitive capacities are limited, Kruglanski proposed that individuals invest more effort in information search when an argument or claim threatens their self-conception or prior beliefs.[38] According to Kruglanski's lay epistemology theory, we search for evidence that supports our preferred view, and stop searching once enough has been gathered. If the model were right, it would be at odds with my claim that motivated cognition relies on two categorically different processes, each with different standards of evaluation. Some evidence does support this view; for instance some

[35] Synder and Cantor, 1979.
[36] Kassarjian, and Cohen, 1965.
[37] Z. Kunda, 1987.
[38] A. W. Kruglanski, 1980; 1990; A. W. Kruglanski & T. Freund, 1983.

studies have shown that individuals scrutinize arguments they disagree with much longer than those that they agree with.[39]

However, there is a large body of evidence that we do deploy two different standards of evaluation, depending on our prior attitudes. Kari Edwards and Edward Smith compared how individuals evaluate arguments concerning issues about which they had strong prior beliefs, including topics like the death penalty and welfare reform. The asked participants to offer a subjective assessment of a single premise argument supporting a conclusion, and then to list reasons why the argument might be correct and reasons why it might be flawed. They found that after controlling for prior knowledge of the subject matter, when individuals disagreed with the argument's conclusions, they i) judged the arguments to be less persuasive, and ii) they listed many more reasons why it might be flawed than correct.[40] In a similar set of studies, Keith Stanovich, Richard West, and Maggie Toplak found that prior beliefs predict the number and kind of evidence individuals are able to retrieve in response to an argument. In a study that asked participants to write down both reasons that favor an argument and reasons that disfavor it, they found that while participants tended to write down the same number of reasons in the time allotted, as in Edwards and Smith's study, prior beliefs determined the ratio of the positive and negative reasons they were able to bring to mind. When participant's disagreed with an argument's conclusion, they listed many more reasons against it than for it. When participant's agreed, the ratio reversed.[41] This work is important because it also shows that confirmation and disconfirmation bias are not artifacts of shallow cognition, and they are not somehow bugaboos of low intelligence. Stanovich, West, and Toplak controlled for intellectual ability using standard measures like IQ tests, and found that there is no relation between intelligence and the tendency to exhibit confirmation and disconfirmation bias. Intellectual gifts do not insulate us from these tendencies.[42]

Indeed, there is some evidence that higher cognitive abilities make people more prone to biased cognition. Recently, experimental political theorists have been exploring motivated cognition to better understand the intransigence of political debate in the United

[39] Neuberg & Fiske, 1987.

[40] K. Edwards & E. Smith, 1996.

[41] K. Stanovich, R. West, & M. Toplak, 2013; see also, K. Stanovich & R. West, 2007, 2008a, and 2008b.

[42] It is possible, however, that intensive training alters the our tendency to deploy these strategies. Academic training, for instance, may inculcate a tendency to deploy the stringent system more often.

States.[43] Dan Kahan, Ellen Peters, Erica Dawson, and Paul Slovic recently conducted a study of how numeracy, an individual's mathematical ability, interacts with motivated cognition. They presented high and low numeracy individuals with the task of evaluating whether or not some statistical arguments supported a conclusion. In a two by two design, they used the very same statistical argument to support a neutral conclusion, that a particular hair-product was effective, and in support of an argument about which participants had prior beliefs, that the death penalty is a deterrent to crime. They found a familiar pattern. Individuals, regardless of numeracy, rate statistical arguments as more persuasive in neutral cases than in cases supporting conclusions the individual disagrees with. What is startling about their result, however, is that they found that high numeracy individuals were far more prone to evaluate the arguments concerning aversive conclusions negatively than were low numeracy individuals. They hypothesize that result is a tradeoff between effort and reward, because thinking is effortful, they propose that high numeracy individuals use their abilities selectively, and tend to apply them when there is a reason to do so, like dismissing an argument they find threatening.[44] Intelligence is no guarantee of objectivity.

There is also evidence that confirmation bias and disconfirmation bias are sensitive to source of information as well as the content. Reactive devaluation is the phenomenon whereby we assess claims differentially based on our preexisting attitude towards the source. Towards sources that we favor, we tend to adopt a permissive epistemic standard, while towards sources we distrust or disfavor, we tend to adopt much more stringent epistemic standards. A classic example is a test conducted by Lee Ross and Constance Stillinger during the mid 1980's. They asked a large sample of people to decide whether a bi-lateral nuclear disarmament agreement between the United States and the Soviet Union would be beneficial to the United States. One group was told that Ronald Reagan had made the proposal, and in that group a very large majority agreed that it would be beneficial. Another group was told that Mikhail Gorbachev made the proposal, and in that group a very large majority concluded that it would not be beneficial. A third group, was asked the same question, but not told anything about the origins of the proposal. That group's conclusion split the difference between the first two, with many concluding it would be beneficial, and many

[43] See John T. Jost, Jack Glaser, Arie Kruglanski, and Frank Sulloway, 2003; Kahan, Slovic, Braman & Gastil, 2006; Lodge & Taber 2013; Sunstein, 2006; Marx, 2007; Westen, Blagov, Harenski, Kilts, & Hamann, 2006; Weber & Stern, 2011;

[44] D. Kahane, E. Peters, E. Dawson, and P. Slovic, 2013.

concluding it would not. This study shows that we evaluate testimony in the same selective manner, depending on how we view the source of that testimony.[45]

5. Debiasing, Counter Hypothesis Search, and Salience Effects

So far I have reviewed the literature on motivated reasoning, and I have argued it supports a number of conclusions. Confirmation and disconfirmation bias represent broad strategies of evaluation, covering self-assessment, the evaluation of arguments and beliefs, memory search, and the evaluation of statistical evidence. These strategies are automatic, and they activate involuntarily depending on an individual's prior attitudes toward a claim or argument. It is automaticity that gives rise to what has been called the illusion of objectivity. We do not feel as if we are biased evaluators because the strategies we apply are not consciously selected. In addition, as shown by the literature on the Wason selection task, strategy selection is intransigent: it is resistant to conscious intervention.[46]

I have not yet drawn a direct connection between motivated cognition and the intuitive appeal of skeptical arguments. I will do so now. Confirmation bias and disconfirmation bias represent broad strategies of epistemic evaluation, and they are the products of what I have called the permissive system and the stringent system. I will now argue that salience effects are a byproduct of one process of the stringent system of epistemic evaluation. I will argue that we find skeptical counter hypotheses in cases of knowledge attribution intuitively appealing because searching for possibilities that would give rise to the same evidence, but which are incompatible with a disfavored focal hypothesis, is one process used by the stringent system in dismissing potential claims. Skeptical challenges feel intuitively compelling because they deploy a strategy we use when seeking to dismiss a disfavored claim. Skeptical challenges feel intuitively misplaced because they are applied to what is typically a favored source of evidence. In what follows I will argue that there is a significant body of empirical research, both from the disconfirmation bias literature, and the debiasing literature, which shows that salience effects in cases of knowledge attribution are a token of a much broader type of evaluation.

As we saw in Chapter 4, salience effects are the product of a process whereby introducing a counter hypothesis undermines our willingness to affirm a focal hypothesis. In

[45] L. Ross & C. Stillinger, 1991. See also L. Ross, 1995; and, A. Ward, M. Katz & L. Ross, 2002.
[46] But see the proviso in footnote 42.

classic presentations of salience effects, a situation is described objectively, and a specific focal hypothesis is described as being actual. Then an unrealized possibility of error is presented to the evaluator of the case, and they are asked whether the agent in the case knows that the focal hypothesis is the case. By focal hypothesis I mean the claim, explanation, or argument that is the explicit focus of evaluation. By counter hypothesis I mean an alternative, incompatible with the instantiation of the focal hypothesis, that would give rise to the same state of affairs. Typically, in salience cases the error possibility is explicitly unrealized. Counter hypotheses need not be explicitly unrealized; they need only be an open possibility that has not already been ruled out.

	Case
Case	John A. Doe is in a furniture store. He is looking at a bright red table under normal lighting conditions. He believes the table is red.
Focal Hypothesis	*The table is red*
Unrealized Error Possibility	However, a white table under red lighting conditions would look *exactly* the same to him, and he has not checked whether the lighting is normal, or whether there might be a red spotlight shining on the table.
Counter Hypothesis	*The table is white with a red light focused on it*
Evaluation	Does John know the table is red?

Table 5.1: Salience Effects and Hypothesis Evaluation

In the case of salience effects, especially given their important role in attempts to diagnose the skeptical problem, the focus has been on knowledge and knowledge attribution. Consequently, as we've seen, the most prominent received accounts have focused on our concept of knowledge as in Fred Dretske's relevant alternatives theory of knowledge, or the semantics of the verb "know," as in contextualists' accounts. I believe that there is strong evidence that salience effects, understood as an intuitive undermining of an individual's willingness to affirm a focal hypothesis, can also be found in studies of i) persuasiveness, ii) subjective assessment of the strength of arguments, iii) spontaneous hypothesis search, and iv) subjective judgments of probability.

5.1 Debiasing: Consider the Opposite and Consider an Alternative Heuristics

Almost as soon as cognitive psychology began to focus on cognitive biases, researchers in the field began to search for means to debias individuals' epistemic evaluations. A survey of the literature reveals that this is remarkably difficult. External rewards like praise and money, and invocations to try harder are quite ineffective at altering individuals' response patterns. One strategy that has consistently been successful is called the "consider the opposite" heuristic.[47] I will argue that it is a misnomer to call this process a debiasing strategy; it is an example of the stringent system at work, and its success as a so-called debiasing strategy is really a result of cueing more stringent standards when permissive standards are already at work. I will also argue that salience effects are instances of this process.

In a survey of the literature on debiasing, Richard Larrick describes the "consider the opposite" heuristic as "nothing more than asking oneself, 'what are some reasons that my initial judgment might be wrong?'."[48] The strategy is widely applicable, and has been used to counter more than just confirmation bias.[49] In one of the first studies to uncover the phenomenon, Charles Lord, Elizabeth Preston, and Mark Lepper investigated motivated reasoning about the death penalty. They first selected participants who had strong prior beliefs, both pro and con, about the efficacy of the death penalty at deterring crime. They then presented both groups with two purported empirical studies to evaluate, one attempting to show that the death penalty is an effective deterrent, and one showing that it is not. In the language I prefer, the conclusion of each study was the focal hypothesis of evaluation. After reading each study, participants were asked to rate how persuasive it was in support of the focal hypothesis (its conclusion). The results showed that most participants were persuaded to some degree, even when the focal hypothesis was at odds with their prior beliefs. However, when subjects were then presented with possible criticisms of the studies—that, for instance, the sample sizes were too small, that the samples were selected in a biased fashion, or that the studies lacked adequate controls—all subjects found these criticisms more persuasive when leveled against the study that conflicted with their prior belief.[50]

---

[47] R. Larrick, 2004: 323-325.

[48] *Ibid,* pg. 323.

[49] See Arkes (1991) and Mussweiler, Strack, & Pfeiffer (2000) for evidence of its efficacy in countering anchoring effects, hindsight bias, and at reducing overconfidence.

[50] C. Lord, E. Preston, & M. Lepper, 1979.

These results support two important conclusions. The initial measure of persuasiveness supports the idea that most individuals try to be even-handed: they don't merely reject objectionable evidence out of hand. Moreover, the differential persuasiveness of the methodological criticisms depending on prior beliefs is indicative of motivated cognition. But most importantly, the preselected methodological criticisms of the focal hypothesis being evaluated amount to counter hypotheses. The participants were asked to evaluate a study that purported to support a focal hypothesis. The methodological criticisms amount to alternative explanations of how the study could have had the results it did, even if the focal hypothesis was false. That subjects found counter hypotheses to a disfavored conclusion more persuasive than counter hypotheses to a favored conclusion suggests they were applying different criteria in evaluation those hypotheses. Perhaps most importantly, Lord, Lepper, and Ross found evidence for what they called a "rebound effect." While subjects were initially persuaded by the studies supporting a focal hypothesis at odds with their own prior beliefs, once they were presented with the counter hypotheses, they reverted to their prior beliefs, sometimes to a position more extreme than they began with.[51] We can see here the pattern of salience effects.

In a follow-up study of the process, Lord, Preston, and Lepper investigated how mentioning an alternative to a focal hypothesis can alter how individuals choose to test a hypothesis. In one experiment, participants were asked to draw up a list of questions to determine whether an unspecified individual was an extrovert. The claim "*X is an extrovert*" functioned as the focal hypothesis. Participants were given two questionnaires, one with questions for identifying the traits of extroverts, and another for identifying introverts. Both groups were advised to be as accurate as possible, and that introverts and extroverts have opposite personality types. The control group exhibited the positive hypothesis testing strategy indicative of confirmation bias, choosing approximately 80% of its questions from the extrovert list, and only 20% from the introvert list. The manipulation in the experimental group was designed to prime them to keep a counter hypothesis in mind, and participants were instructed to remember that the individual they'd be testing might not be an extrovert. The priming was effective at altering their test strategy: while the group still leaned toward confirmation bias by choosing a positive test strategy 56% of the time, the

[51] *Ibid,* pg: 2105.

invocation to consider the opposite doubled the rate at which they used a negative test strategy to 44%.[52]

Lord, Preston, and Lepper clearly see their results as the product of a "corrective strategy," a heuristic that can help overcome confirmation bias, and not, as I claim, as priming a broad set of processes with a different function. What they don't show, and what the tradition that followed them could not show either, is that the "consider the opposite" strategy somehow brings performance on reasoning tasks into line with some normative standard of inquiry. There is no consensus, even in the sciences, about whether empirical inquiry aims at confirmation or disconfirmation. Instead, what they find is that the strategy *alters* the strategy individuals use to search for and test hypotheses.[53]

In a similar study Asher Koriat, Sarah Lichtenstein, and Baruch Fischhoff investigated the role of the "consider the opposite" strategy in altering confidence judgments. The explicit goal was to uncover myside bias in confidence assessments. Their results are, in fact, a further example of the role of salience effects in motivated reasoning. They asked participants to answer a number of general knowledge questions. They then asked subjects to indicate their confidence in those responses. The authors then broke subjects into two groups, asking one group, to provide reasons supporting their previous responses and the other group to provide reasons contradicting their previous responses. Each group was then asked to rate their confidence in their responses again. The authors found that those in the first group did not significantly alter their confidence assessments, but that those in the second did.[54] The study demonstrates salience effects with confidence judgments. Prompting individuals to generate counter hypotheses to a previously given focal hypothesis undermines confidence in the focal hypothesis.

---

[52] C. Lord, E. Preston, & M. Lepper, 1984: 1238.

[53] It's worth noting here that the heuristics and biases literature has a misleading name. The term "bias" suggests that the kinds of epistemic processes that I've discussed here are local aberrations from a more general, objective process of evaluation. Bias are sometimes called cognitive blind spots, or cognitive illusions, and metaphors like these liken biases to a small area where perception fails to function as it should and normally does. This implication is inaccurate. The accumulated evidence that I have canvassed here, and the even greater wealth of empirical work that space considerations forced me to forego discussing, overwhelmingly suggests that most cognition is motivated cognition. Consciously objective reasoning, where evaluative strategies are uniformly applied irrespective of the evaluator's prior attitudes and information about a target claim, is the exception, not the norm. It may, sometimes, be difficult to recognize this because philosophical training, and indeed all academic training, aims toward fostering just this approach. Indeed, whether or not academic training *succeeds* in cultivating this stance is an open question; the rhetoric and self-conception most certainly embody this ideal.

[54] A. Koriat, S. Lichtenstein, & B. Fischhoff, 1980.

The strongest evidence for my claim that salience effects are the product of a general strategy of epistemic evaluation comes from a series of studies conducted by Edward Hirt and Keith Markman. Their goal was to investigate how the "consider the opposite" heuristic worked, and if it functioned by comparison of two mutually exclusive hypotheses (i.e., that the death penalty does / does not deter crime), or if it functioned by comparing a focal hypothesis to any alternative (i.e., that the Expos would win the pennant, or that the Pirates, Reds, Cardinals, or Marlins would win). Thus, they hypothesized that considering an alternative to a focal hypothesis, not just the focal hypothesis' falsity, would undermine epistemic assessment. Secondarily, they wanted to investigate the power of the heuristic, and whether considering some alternatives undermine epistemic evaluations more than others.

Their experiments were based on a type of confirmation bias known as explanation bias. Explanation bias is the effect whereby explaining how a focal hypothesis might be possible causes us to judge that hypothesis to be more likely than when asked to judge its likelihood without first offering an explanation.[55] In an experiment conducted in 1993, Hirt and Markman asked participants to read about five Major League Baseball teams in the National League East prior to the start of the 1993 season. At the time the Montreal Expos were the considered pick to win the pennant, the Cardinals and Cubs were reasonable choices, and the Reds and the Marlins (an expansion team new that year) were very implausible choices. In the control condition, participants were asked to explain a divisional win by the Expos, and then render a judgment of its likelihood. In the experimental conditions, the participants were also asked to explain a division win by the Expos, and render a judgment about its likelihood. But rather than simply ask them to consider the opposite, i.e., that the Expos don't win the division, in each of four other conditions they asked participants to explain a divisional win by another team (either the Reds, Cardinals, Cubs, or Marlins) and also judge the likelihood of this event. They found that individuals that considered an alternative to the Expos winning judged it to be less likely than those that did not consider an alternative. Critically for our understanding of salience effects, and the empirical results I report in Chapter 4, they found that the effect was much more powerful when considering plausible alternatives, like the Cardinals, than when considering

[55] C. Lord, E. Preston, & M. Lepper, 1984

implausible alternatives like the Marlins.[56] On the basis of a series of similar studies, Hirt and Markman propose to rename the "consider the opposite" heuristic the "consider an alternative" heuristic.

This study, along with the others reported in the same paper, clearly have the structure I proposed above. A focal hypothesis, that the Expos will win, is evaluated and then a counter hypothesis is introduced, that one of the four other teams in the division will win, which undermines participants' judgments about the subjective likelihood of the focal hypothesis.[57]

5.2 The Psychology of Salience Effects

As I believe the evidence shows, salience effects on knowledge attribution are a special case of what the cognitive psychology literature has come to know as the "consider the opposite," or "consider an alternative," heuristic.[58] Salience effects have a basic structure: all things being equal, epistemic evaluations of a given focal hypothesis will tend to be high. When a counter hypothesis is introduced, where a counter hypothesis is an alternative account of the evidence for the focal hypothesis, epistemic evaluations tend to decrease. This pattern is not confined to knowledge attributions, but appears across the

---

[56] E. Hirt & K. Markman, 1995: 1082.

[57] Hirt and Markman's work is by no means unique; it is merely the closest example in the literature on cognitive psychology to what epistemologists describe as salience effects. Neither is the work canvassed in this section an exhaustive account of the power of considering alternate hypotheses in altering epistemic assessment. See also: Sherman, S. J., Cialdini, R. B., Schwartzman, D. F , & Reynolds, K. D. (1985). "Imagining can heighten or lower the perceived likelihood of contracting a disease: The mediating effect of ease of imagery." *Personality and Social Psychology Bulletin*, 11, 118-127; Kahneman, D., & Miller, D. T. (1986). "Norm theory: Comparing reality to its alternatives." *Psychological Review*, 93, 136-153; Wells, G. L., & Gavanski, I. (1989). "Mental simulation of causality." *Journal of Personality and Social Psychology*, 56, 161-169.

[58] In an experimental study designed to demonstrate salience effects with likelihood judgments, I recruited 101 individuals on Amazon Mechanical Turk (Age *M:* 30.1; Female: 34%; Education: 35% bachelors or higher; Race: 79% White or European) who were compensated at minimum wage. The study involved two independent conditions. In the control condition participants read the following story, and rated the likelihood that the Oakland A's would win the pennant on a ten point likert scale, where 1 = very unlikely and 10= very likely: *Right now the Oakland Athletics are the best team in baseball. They have won 59 games and only lost 36. Their pitching staff is one of the best in baseball, and it has given them a clear advantage over other teams. History has shown that teams with a comparable record at this point in the season have a very high chance of winning their division.* In the salience condition, the following unrealized possibility was mentioned in addition to the text above: *However, while pitching is the Athletics' strength, pitchers are the players most prone to injury, and the hardest players to replace. If one or more of the Athletics' pitchers were to suffer an injury between now and the end of the season, their chances of winning the division would be greatly reduced.* There was a statistically significant difference, $t(99)=3.87$; $p < .0001$; $d = 0.75$ (a "large" effect) between likelihood judgments in the control condition ($n=44$, $M=8.34$, $SD=1.7$) and the salience condition ($n=57$, $M=7.07$, $SD=1.59$). Note: subsequent to this study, the A's, in fact, went on to loose the division after a historic September collapse, and were eliminated from the playoffs in a wild-card game with the Royals.

literature in debiasing. The mistake of this literature was to think that epistemic devaluation is somehow a process that leads to more objective processing of information. There is no evidence for this. Instead, it is simply the expression of a different kind of bias, one that structurally favors dismissing a position rather than accepting it.

To sum up the findings, we see that a number of kinds of epistemic assessments exhibit a pattern very similar to salience effects. Considering the opposite (Lord, Preston, and Leppert, 1979) of a focal hypothesis decreases subjective judgments of persuasiveness. Being told to consider the opposite of a focal hypothesis prompts negative hypothesis testing, and causes people to search for evidence for a counter hypothesis (Lord, Preston, and Lepper, 1984). Being asked to spontaneously generate counter hypotheses undermines confidence judgments (Koriat, Lichtenstein, & Fischhoff, 1980). Being confronted with an unfavorable conclusion automatically prompts individuals to search for counter hypotheses (Edwards & Smith, 1996). Being asked to consider an alternative counter hypothesis to a focal hypothesis undermines subjective likelihood judgments (Hirt & Markman, 1995). Finally, the subjective plausibility of a counter hypothesis is inversely proportional to its power to undermine subjective likelihood judgments (also Hirt & Markman, 1995). As I will argue in the following section, the fact that we find this pattern across modes of epistemic evaluation is strong evidence against the linguistic contextualist's account of salience effects. But, before we can compare psychological contextualism to linguistic contextualism, we must first examine how and why we possess these two systems of epistemic evaluation, and why they seem to aim to produce conflicting evaluations of claims and arguments.

6. Context and Epistemic System Activation

Why would we have, as I propose, two systems of epistemic evaluation, each employing incompatible standards, one aimed at confirming beliefs and another aimed at dismissing them? The conflict between the standards would seem to create the potential for the incoherence of our epistemic endeavors: sometimes we would affirm a claim, and other times dismiss it, depending on which system activates.

According to epistemic contextualism's theory of the semantics of "know," we have a single system of evaluation, with a single sliding standard of evaluation, that becomes more

or less stringent depending on contextual factors.[59] According to the view I favor, we have two categorically different standards of evaluation, one permissive and fallibilist, and another more stringent and infallibilist. According to the linguistic contextualist, contextual factors including stakes and whether an error possibility has been introduced determine the stringency of the standards of evaluation.[60] According to the view I propose, the permissive system and the stringent system are automatically activated by a different set of contextual factors. The content of the claim or argument being evaluated, and the source of the claim or argument being evaluated determine which standard of evaluation is activated. The permissive system is the default strategy of evaluation, and it is activated when the claim or argument to be evaluated is one toward which we have a neutral or positive prior attitude, or when the source of the claim is one we trust. The stringent system is activated when the claim or argument to be evaluated is one toward which we have a negative prior attitude, or comes from a source we distrust.

According to the received view of the function of human reasoning, reasoning is individualistic and aims at producing true, useful beliefs. Though not incompatible with other views, the linguistic contextualists' account of the semantics of "know" is of a piece with this view of reasoning. It posits a single standard for attributing knowledge, and only elaborates the view by adding that contextual factors can induce us to expend extra effort in making evaluations. On the assumption of this view, the literature on biased cognition I've just reviewed represents a selective deviation from reasoning's intended function. But, the consideration works in the other direction, too. The breadth of the literature I've just canvassed suggests the received view is descriptively inadequate. Dan Sperber and Hugo Mercier have argued that the function of human reasoning is social, not individualistic, and that our rational capacities evolved to aid in arguing with our epistemic peers, not primarily as part of the solitary search for information.[61] This argumentative theory of reasoning can help answer the question I began the section with. We have two systems of evaluation because they both function to help produce and evaluate arguments and claims. The function of the permissive system is to produce arguments and help defend claims toward which we have positive prior attitudes. The function of the stringent system is to produce

[59] K. DeRose, 1995.
[60] K. DeRose, 2009: Chapter 1.
[61] D. Sperber & H. Mercier, 2011.

arguments against rivals and to evaluate claims toward which we have negative prior attitudes.

Sperber and Mercier's account of reasoning emphasizes the importance of collaborative communication throughout evolutionary history, especially to facilitate tasks like hunting, gathering, raising children, and coordinating collective projects. The problem of communication is that individuals must be able to discriminate reliable and trustworthy information from unreliable and deceptive information. Without this ability an individual would be easy to take advantage of (here are some magic beans). Sperber and Mercier propose that reasoning is fundamentally argumentative because its function is to promote epistemic vigilance. Arguments, both for a claim and with another individual, are a way to evaluate potential claims and police communication for deception. They defend the view by pointing to a number of predictions it makes, and for which there is strong evidence: that i) we are better at reasoning in argumentative contexts than in the abstract, that ii) we reason better in groups than alone, that iii) argument production exhibits myside or confirmation bias, that iv) confirmation bias will be stronger when reasoning alone than in groups, and v) that we will choose actions that are easier to justify.[62]

I do not want to attempt to evaluate this view; the considerations for and against it are too numerous to sort out here. Be that as it may, the argumentative theory of reasoning is a descriptive account of the evolutionary development of reasoning, not a normative account of how it should function. And with that in mind it can help illuminate the proposal I make. We have a permissive system of evaluation to make arguments and evaluate the claims of those we trust. We have a stringent system of evaluation to evaluate the claims of those we distrust. It makes evolutionary sense that the former would be biased toward confirmation. Cognition consumes time and effort, and we should not expend as much effort to evaluate the claims of those that we already trust, or claims that cohere with our prior commitments. On the flip side, when confronted with a claim that conflicts with our prior commitments, especially coming from a source that we distrust or think might intend to deceive us, it makes evolutionary sense to both expend more effort in evaluating that claim, and that the process should be biased toward dismissal. We should not ignore what our rivals say, but we should be willing to accept it only if we can find no alternative.

[62] *Ibid*, pgs: 61-71.

This helps explain the context sensitivity I posit. The selection of the strategy of evaluation is determined by contextual factors: the content of the claim or argument and its source. It also helps respond to a possible objection. Others have proposed that concept of knowledge is inconsistent.[63] But these proposals have been criticized because they would seem to inevitably lead to incoherence in everyday life: sometimes we would know a claim *P*, and other times we would not. The criticism is that while these views can help us make sense of some competing intuitions, like those that sometimes draw us toward fallibilist theories of knowledge as in Gettier cases, and those that draw us toward infallibilist theories of knowledge, as in Dretske's zebra and duck salience cases,[64] they do not offer enough guidance about when those standards will be active to avoid pervasively incoherent attitudes. My proposal does.

I am not offering a normative account of knowledge here. I am offering a description of our epistemic psychology. Many people think or psychology is kludgey for evolutionary reasons.[65] Those reasons apply equally well here. An inconsistent set of procedures could be an ecological problem, leading to all sorts of inconsistent actions if the cues that trigger the systems overlapped. However, the evidence surveyed in the previous section suggests that the cues do not overlap, and that they are in fact well delineated.

Typically, accepting, rejecting or having a neutral stance toward a claim or argument would exhaust the possibilities, and consequently there would be no practical conflicts in our evaluations. Conflict could still arise, but the circumstances would be limited. One set of cases would be where either i) where both standards were activated in evaluating the same claim or argument, or ii) where both standards where activated toward a single source. We've seen evidence of this already in the literature on the "consider the alternative" strategy. The results suggest that we tend to settle for the results of the strategy we apply last. The "rebound effect" observed by Lord, Preston, and Lepper shows that whichever strategy is applied last tends to moderate the earlier evaluation: when a confirmatory consideration is considered after counter evidence, epistemic evaluations move in the direction of confirmation; when a disconfirmatory consideration or counter hypothesis is

[63] M. Weiner, 2009.
[64] See Chapter 4.
[65] S. Stich, 2006; G. Marcus, 2009.

considered after a confirmatory consideration, evaluations move in the direction of dismissal.[66] There is no practical incoherence because we apply the strategies sequentially.

However, one circumstance where incoherence can arise is when one standard is applied to a claim, and then the other standard is applied to another claim or argument that is inconsistent with the first. Because the evaluations run in parallel, and because we do want consistent beliefs, conflict is possible. This conflicting parallel evaluation is the conflict at the heart of the skeptical problem.

Another important objection deserves consideration. There is a reason that, as Trent Dougherty and Patrick Rysiew claim, fallibilism has "near-universal acceptance," and it is that infallibilism makes knowledge too hard.[67] Wouldn't the stringent system I propose be far too indiscriminant in rejecting disfavored claims? Perhaps, so the objection goes, it is ecologically rational to subject claims we reject or claims from rivals to more stringent scrutiny, but this proposal would amount to automatic rejection. Classical infallibilism requires that we rule out *all* possible alternatives before we can know that *P*.[68] As radical skepticism shows, it is always possible to dream up new counter hypotheses, and so it is impossible to know *P*, even when it is a quotidian claim like "I have hands." But, if the standard the stringent system applies to the claims of rivals is classic epistemic infallibilism, then we would end up rejecting every belief we thought to be false, and every claim or argument made by a rival. For, all we would need to do is find some possible counter hypothesis to the claim being evaluated.

This objection turns on the idea of classic, high-standards infallibilism. I believe there is reason to believe that the standard used by the stringent system is to rule out merely plausible alternatives, not all possible alternatives.[69] On this account we must accept a claim P if there are no plausible counter hypotheses to it. This standard is much lower than that of classical infallibilism. The evidence for this moderate infallibilism comes from the salience studies in Chapter 4, and the experiments of Hirt and Markman above. In experiments 2a-d we found that when a counter hypothesis is subjectively plausible, it undermines knowledge

[66] C. Lord, E. Preston, & M. Lepper, 1979: 2105

[67] T. Dougherty & P. Rysiew, 2009: 123.

[68] B. Reed, 2012.

[69] Compare this with relevant alternatives theory. That theory was designed to account for salience effects cause by both plausible counter hypotheses and implausible skeptical counter hypotheses. It floundered because of the difficulty of refining "relevant" in a way that was more informative than just "whatever the evaluator takes to be relevant."

attribution. As counter hypotheses become more subjectively implausible, their ability to undermine knowledge attributions diminishes. When a counter hypothesis is highly implausible, it doesn't alter knowledge attribution at all - it is as if a counter hypothesis was not mentioned. Hirt and Markman's "consider the alternatives" results above mirror the studies I conducted: participants' judgments of the likelihood of a focal hypothesis were undermined in inverse proportion to the likelihood of the alternative they were asked to consider.

Rejecting a disfavored focal hypothesis does not involve accepting the counter hypothesis. The attention checks in experiments 2a-d asked participants whether the focal hypothesis was true, or whether the counter hypothesis was true. The vast majority of participants (96%) recognized that the focal hypothesis was true, and analyses of knowledge attribution were confined to their responses. The stringent system asks *must I believe P?*, and if it can generate a plausible counter hypothesis, the answer is *no*, and the evaluator will dismiss the claim of a rival. This strategy is ecologically rational. If an individual is confronted by a disfavored claim, and they cannot think of a counter hypothesis that is either true or at least possible, then they should not dismiss the claim, and give it further attention. This is not a procedure that aims at the truth; it aims at economizing on effortful thought while avoiding adaptively dangerous errors. How plausible must a counter hypothesis be to count as a threat to a claim? I don't have an account of this yet, and it deserves empirical attention. It is presumably a complicated function that depends on the content of the counter hypothesis and the evaluator's background knowledge.

7. The Psychological Diagnosis of the Skeptical Problem

I have now assembled and previewed the diagnosis of the skeptical problem that I favor. The skeptical problem arises from the joint inconsistency of the ordinary intuition, the Cartesian intuition, and the intuition of conflict. The appeal of this inconsistent triad is a byproduct of applying different standards of evaluation toward quotidian perceptual claims, and to arguments like AI that support the conclusion that those perceptual claims are mistaken.

We have two epistemic systems, a permissive system that gives rise to confirmation bias, and a stringent system that gives rise to disconfirmation bias. Confirmation bias invokes a fallibilist standard whereby epistemic evaluation is sensitive to evidence that

confirms a position. We ask *can I believe P?* and answer yes when there is some justification for P. Disconfirmation bias invokes an infallibilist standard whereby epistemic evaluation is sensitive to the possibility of plausible alternatives to a position. We ask *must I believe P?* and answer no whenever there is a plausible counter hypothesis.

The evidence from the motivated reasoning literature shows that the activation of the permissive and stringent standards is context-sensitive and involuntary. They differentially activate depending on the content, or source, of the focal hypothesis under evaluation and the prior beliefs of the evaluator.

The evidence from the debiasing literature shows that when an individual finds a focal hypothesis unfavorable they i) spontaneously generate more counter hypotheses than when a focal hypothesis is found favorable, and ii) that considering counter hypotheses undermines a) knowledge attribution, b) judgments of persuasiveness, c) judgments of confidence, and d) judgments of likelihood.

I propose that disconfirmation bias involves the following two-stage process: unfavorable focal hypotheses are evaluated against an infallibilist standard, and when considering such a hypothesis people search for plausible counter hypotheses. The more and more plausible counter hypotheses they generate, the lower their epistemic assessment of the focal hypothesis.

Towards favored sources of information, like perception and the testimony of friends, and toward neutral subjects about which we do not have strong prior convictions, we tacitly deploy the permissive system for evaluating claims. Take the claim "I have hands" as an example. Because it is a claim that coheres with our past evidence and for which there is ample perceptual evidence, and because it does not have an untrustworthy source, the permissive system is tacitly activated. It asks, in effect, *can I know that I have hands?* Because the permissive system searches for confirming evidence, and because there is ample evidence to confirm that we have hands, we intuitively answer *yes*. The permissive system underwrites the ordinary intuition.

Towards disfavored sources of information, like the arguments of rivals, and towards claims that threaten to contradict strong prior convictions, we tacitly deploy the stringent system of evaluating claims. The stringent system applies an infallibilist standard, and does so in some cases by searching for counter hypotheses. Arguments like AI take advantage of this. Typically we do not search for counter hypotheses when considering favored claims.

Arguments like AI appeal to the stringent system's standard by introducing a counter hypothesis that is incompatible with the focal hypothesis. By introducing one, it prompts us involuntarily to ask, in effect, *must I believe that I have hands?* Typically, the stringent system attempts to answer the question by searching for alternatives: we must accept a claim when there is no plausible alternative to it, and we need not accept it when there is a plausible alternative. Arguments like AI preempt the need for such a search, and activate it directly. Because being a brain in a vat is at least possible, and we cannot rule out, we intuitively answer the question *no*.[70]

What about other associated claims? Why don't we stringently evaluate the claim *I am a handless brain-in-a-vat?* Wouldn't stringently evaluating this claim, especially since it threatens a favored belief, lead us away from the skeptical problem? I think we do stringently evaluate it, and that there is a plausible alternative to it. I think the positive claim fails against the standards of both the stringent and the permissive system. But the skeptic isn't trying to convince you that you are a brain-in-a-vat, rather she is challenging you to show that it is false.

Must I believe that *it's possible that I'm a handless brain-in-a-vat?* According to the proposal I outline, we search for plausible alternatives, and we do not dismiss a claim unless we can find one. On the account I give we are only compelled to accept a claim if there are no plausible alternatives. But, it seems like we must accept that *it is possible that I'm a handless brain-in-a-vat* even though there are plausible alternatives. It would appear that my proposal has difficulty handling alethic evaluations. I believe alethic judgments about necessity and possibility are derivative upon the hypothesis search procedure I outline, and not subject to it. We make judgments about what is possible and what is necessary by conducting a hypothesis search.

Arguments like AI above formalize the structure of the evaluations we find in salience effects cases. When we ask *must I believe P?*, we answer *no* when there is an alternative that is consistent with the evidence for P. Must we believe a study showing the death penalty is an effective deterrent? Not if there is an alternative explanation of the study's results - perhaps that it had a biased sample. It's critical, however, that a counter hypothesis is an alternative that is not already ruled out by the available evidence or previous

[70] I'll address the implausibility of radical skeptical scenarios shortly.

knowledge. If we think that they are true, then so much the better, but we must at least not think that they are false. Consider the following example,

> John A. Doe is in a furniture store. He is looking at a bright red table under normal lighting conditions. He believes the table is red.
>
> However, a white table under red lighting conditions would look *exactly* the same to him.
>
> John worries that the lighting might be abnormal, and he checks closely to make sure that the lighting is natural and not misleading.
>
> *Does John know it is not a white table under red lighting?*
>
> *Does John know the table is red?*

In this case a counter hypothesis is introduced, but it is ruled out by the agent's investigation. Does he know that it is not a white table under red lighting? I think the intuitive answer is *yes*. Likewise, does he know the table is red? Again, I think the intuitive answer is *yes*. In recent study, Joshua Alexander, Chad Gonnerman, and I tested screened-off salience cases such as these, and found that participants attributed knowledge at the same high level as when a counter hypothesis is not made salient.[71] We search for counter hypotheses that are compatible with our available evidence and knowledge, and so they must possibilities that we do not think to be false. It is this feature of the nature of the search for counter hypotheses that explains the attraction of arguments like AI.

	Standard Salience Case	Cartesian Argument from Ignorance
0.	John knows that table is red	John knows he has hands
1.	John doesn't know that the table is not white with a red light shining on it.	John doesn't know he's not a brain-in-a-vat
2.	If John doesn't know that the table is not white with a red light shining on it, then he doesn't know that it is red.	If John doesn't know he's not a brain-in-a-vat, then he doesn't know that he has hands.
3.	Therefore, John doesn't know that the table is red.	Therefore, John doesn't know that he has hands

Table 5.2: The Argument from Ignorance and Stringent Epistemic Evaluation

[71] J. Alexander, C. Gonnerman, & J. Waterman, 2014.

The permissive system explains the intuitive appeal of 0. The intuitive appeal of 1 is a byproduct of the nature of the stringent system's search procedure. The stringent system evaluates claims by searching for counter hypotheses that are true, or at least plausible, but not thought to be false. So, a counter hypothesis must be one that an evaluator does not know to be false. Counter hypotheses in salience cases are by design alternatives about which there is no evidence to rule them out, and radical skeptical hypotheses are likewise by design possibilities about which there is no evidence to rule them out. Consequently, they are possibilities we do not know to be not the case, and claims to this effect strike us as intuitive. The bridge premise, 2., strikes us as intuitive because well-chosen counter hypotheses are just those claims that imply the falsity of the focal premise. Arguments with the form of AI are valid, and so the conclusion, 3., follows as a matter of course.

This account can also explain why dogmatic skeptical challenges fail. A dogmatic skeptic can simply assert "you might be mistaken" as a challenge to a focal hypothesis like "I have hands." But this kind of challenge seems far less threatening than classical versions of AI. A generic counter hypothesis like "you might be mistaken" is possible in almost any circumstance, but it is very weak as an alternative to our beliefs about having hands because it is very implausible that we would be so mistaken. What the

Skeptical arguments like AI induce us to apply the stringent system's standards of evaluation to favored hypotheses by introducing a possible counter hypothesis, and then working out its implications. As I argued above, the stringent system limits its searches to plausible counter hypotheses, and in Chapter 4 I gave evidence that we are typically indifferent to implausible counter hypotheses. Radical skeptical hypotheses about alien super-psychologists and malevolent demons, are, however, quite implausible. Thus, while the AI formula works in the way I've sketched above for *plausible* hypotheses, the diagnosis of radical formulations of AI involves one more step. The skeptic must prime us to overcome our indifference to the implausible. Experiments 3a and 3b in Chapter 4 show us this is possible trough what I called Cartesian priming. By sequentially presenting counter hypotheses to a focal hypothesis from most plausible to least plausible, the skeptic can overcome indifference to implausible defeaters. We saw in these experiments (see figure 8., Chapter 4) that this effect is an all-or-nothing affair. Cartesian priming does not affect all individuals. In plausible error cases (experiments 2a-d) responses were unimodal and most participants did not attribute knowledge to the agent in the vignettes (*M knowledge*

attribution=2.95 - 3.43).[72] In very implausible error cases responses were unimodal and most participants did attribute knowledge (*M knowledge attribution*=4.22 - 4.34). In very implausible error cases after Cartesian priming, responses were bimodal, with approximately half the participant pool attributed knowledge at a very low level (*M knowledge attribution*=1.94) and half the participant pool attributed knowledge at a very high level (*M knowledge attribution*=4.8). Thus, one reason skeptical arguments like AI only affect some people is that Cartesian priming only affects some people.

This leaves open the question of *why* Cartesian priming is an effective means of overcoming our indifference to implausible counter hypotheses. I'll return to this question momentarily, but first it is now time to compare the psychological diagnosis of the skeptical problem to linguistic contextualism.

8. Psychological Contextualism and Linguistic Contextualism

In Section 3 of this chapter I laid out a set of desiderata for a diagnosis of the skeptical problem. They included: i) empirical adequacy, ii) theoretical unification, and iii) charity. These desiderata offer a way for us to measure the virtues of these rival accounts. Throughout the last three chapters I have made a case about each of these points individually, and so I will not rehearse those arguments here. Instead I will summarize those conclusions.

8.1 Empirical Adequacy

Chapter 3 empirically investigated the intuition of conflict. A key result was that there are individual differences in whether someone evaluating AI will experience the conflict. Chapter 4 empirically investigated salience effects. Those studies had two key results. The first was that the size of salience effects is inversely proportional to the plausibility of the counter hypothesis prompting it. The second was that Cartesian priming overcomes indifference to the implausible, but only for some people.

Linguistic contextualism has difficulty explaining either of these facts. One of the leading objections to linguistic contextualism is that it must posit an implausible kind of error theory to explain the intuition of conflict. Linguistic contextualists claim we are prone

[72] Higher number represent attribution of knowledge, lower numbers denial; the midpoint of the scale was 3.5.

to systematic semantic blindness, but as the objection goes, we do not observe any similar kind of systematic misunderstanding of semantic facts elsewhere in linguistics. I argued in Chapter 3 that cognitive psychological accounts can much more easily explain individual differences in reasoning like this. The cognitive psychology literature and the heuristics and biases tradition are ripe with examples of individual differences like this. I believe there is much more to be said about the intuition of conflict, both experimentally and theoretically, but on the whole cognitive accounts do better.

Linguistic contextualism also has difficulty explaining the observed behavior of salience effects. Linguistic contextualists tend to claim that introducing counter hypotheses generate salience effects by raising the standards for knowledge attribution. This in itself is a somewhat puzzling claim. According to Keith DeRose, who has the most sophisticated account of how this occurs, mentioning a possibility of error prompts us to apply greater scrutiny to a focal hypothesis. We do this by searching through relevant possible worlds and examining whether our evidence for the focal hypothesis is strong enough to rule out cases where it is false. On this view, "as standards go knowledge go up, the sphere of relevant world becomes larger," and so we must examine worlds where the counterhypothesis is true, and see if we can rule it out.[73] But because radical skeptical hypotheses cannot be ruled out, we accept that we do not know them to be false. What is puzzling about this account is that it is not clear that examining just any counter hypothesis really is applying higher standards for knowledge. Certainly examining many alternatives to a focal hypothesis is more stringent than examining just one, but it's not clear that considering implausibly far-fetched hypotheses amounts to higher standard investigation. If I have completed a long, arduous mathematical proof, checking all the steps for, say, mistakes in addition counts as a more stringent test of its validity than not checking for mistakes. But does considering the possibility that I am a brain-in-a-vat, or that I am drugged, or that my mathematical education was an elaborate ruse count as a more stringent test of its validity than checking it for mistakes in addition? Not clearly.[74] Perhaps if introducing one counter hypothesis, however plausible, induced us to check many counter hypotheses, then it would come closer to actually introducing higher standards. The evidence from the salience studies in Chapter 4 suggests that this is not the case: participants took only marginally longer to answer the

[73] K. DeRose, 1995: 37.

[74] Michael Williams made this point in a personal communication.

target question in cases where a counter hypothesis is mentioned than in those where one is not mentioned. It would seem that we only consider the mentioned counter hypothesis, and we do not conduct a broader search.

Linguistic contextualism does not predict either indifference to the implausible or Cartesian priming. The former is a serious shortcoming. The theory is designed to explain the appeal of arguments like AI, and it explicitly assimilates radical skeptical hypotheses to the more mundane kind of error possibilities that give rise to salience effects. The failure to explain how skeptical hypotheses become intuitively compelling is a major shortcoming. I have no doubt that it will be possible to amend the theory to accommodate both these empirical findings. But be that as it may, they will be accommodations and not predictions of the central empirical claim of contextualists, namely that "know" is a context-sensitive term with semantics similar to scalar absolute adjectives. Accommodating these features will be *ad hoc*.

How does psychological contextualism fare? Rather well, as it should, seeing as I expressly conceived it as an explanation of the empirical results reported in Chapters 3 and 4. It could be argued that this leaves the theory no better off than linguistic contextualism. Psychological contextualism was in fact tuned to fit the data, while linguistic contextualism must be retuned to fit the data. This is where auxiliary factors come to the fore.

8.2 Theoretical Unification

We should prefer theories that are theoretically unified, and cohere with a broad set of results, to those that are less unified and do not cohere with a broad set of results. In this chapter I have argued that salience effects are a byproduct of disconfirmation bias. I have reviewed results showing that unfavorable claims automatically prompt searches for counter hypotheses. I have also argued that salience effects are not restricted to knowledge attribution, but can also be found in epistemic evaluations of persuasiveness, judgments of likelihood, and judgments of an argument's strength. The account of the psychological origins of salience effects in knowledge attribution that I have proposed integrates a very large body of research in cognitive psychology.

Linguistic contextualism's account of "know" as having semantics similar to scalar absolute adjectives is at odds with some more general claims in linguistic theory, and with the empirical results I present here. Chapter 3 reviewed some of the many linguistic tests for

context-sensitive terms, and found that "know" does not pass them. In a direct test of the analogy between "know" and scalar absolute adjectives like "flat," experiment 2 in Chapter 3 showed that while it is possible to use arguments like AI to induce "flatness skepticism," ordinary speakers do not experience the intuition of conflict with terms like "flat." By contrast "know" behaves differently, some individuals do experience the intuition of conflict, while others do not. While this does not show that "know" is not context-sensitive in the way claimed by contextualists, it does show that the intuition of conflict is not a product of a more general semantic phenomenon. The more ways in which contextualists must claim that "know" has *sui generis* semantic features to accommodate the empirical evidence, the less the are able to achieve theoretical unification. Psychological contextualism does not offer a positive account of individual difference in experiencing the intuition of conflict, but neither is at odds with a wide body of theory. There are many individual differences in the heuristics and biases literature.

Perhaps a natural reply for linguistic contextualists at this point is to posit pervasive epistemic contextualism. As we saw, support for psychological contextualism can be found in the presence of salience effects in other forms of epistemic evaluation. Perhaps linguistic contextualists could argue that if we find evidence for context-sensitivity with "know," then we should also expect to find it with other terms of epistemic evaluation, like "persuasive" and "likely." In this way linguistic contextualism would predict the same results in the studies I appeal to in support of the psychological account. The results from studies of hypothesis search offer one reason to resist this. We saw in the work of Kari Edwards and Edward Smith, and in separate studies by Charles Lord, Elizabeth Preston, and Mark Lepper that when people are confronted with unfavorable focal hypotheses they spontaneously search for plausible counter hypotheses.[75] This is evidence for categorically different testing strategies, not an incremental change in standards or similar form of pragmatic accommodation. People do not merely apply incrementally more effort in their search, as one might expect if standards are to be measured on a scale, they instead conduct a different kind of search.

This too could be accommodated by a more pervasive kind of linguistic contextualism, but now the contextualist would have to posit a different kind of context-

[75] C. Lord, E. Preston, & M. Lepper, 1984; Kari Edwards and Edward Smith, 1996.

sensitivity. To accommodate this data they would have to posit that epistemic terms have multiple, categorically distinct meanings; in other words, they might claim that epistemic terms are ambiguous or polysemous, not scalar. Such a move might compound the challenge they face in positing an error theory to explain why we fail to recognize this: we fail to recognize this with "know" *and* with other epistemic terms of evaluation.

My point is not that there is no linguistic model that contextualists could adopt to explain the observed pattern of behavior. One view of ambiguity, polysemy, and vagueness is that they are not distinct linguistic phenomena, but examples on a continuum. Ambiguity is the most easily tested, by the intuitive sense of crossed readings of the term. Sentences like "Tom went to the bank (financial institution) and so did Bill (river's edge)," can only be said facetiously, or as a pun, and reveal that "bank" has distinct senses. The term "aunt" has distinct meanings (father's sister and mother's sister), but this is hard to detect by way of crossed readings. The intuitive ordinariness of "Tom has an aunt (mother's sister) and so does Bill (father's sister)" reveals that "aunt" is vague. Polysemy lies somewhere in between, the sentence "Tom played (at checkers) Bill in the morning and Hamlet in the afternoon (as an actor)," admits a normal interpretation (Tom played two people at checkers) and a facetious one (Tom had a game of checkers and was in a play).[76] If epistemic terms are to be placed somewhere along this continuum, and we attempt to explain the failure to recognize this in terms of an error theory, the size of the error will depend on the model that is adopted. Models positing ambiguity will involve larger error theories, and models using vagueness smaller ones.

To reiterate, my point is not that "know" is not context-sensitive. There are surely models that can accommodate the data. I have only argued against one specific model. My point is, instead, that the semantics of "know," whatever they are, are a partial byproduct of how our epistemic psychology drives applications of the term. It is yet a further point to claim, as I do, that appeal of AI does not depend ultimately on semantic facts. The psychological model I favor helps explain the behavior of epistemic terms in linguistic tests. The permissive system and the stringent system activate automatically in response to the nature of the content at issue and its source. Once one is active, it stays active. Typically we default to the permissive system when the content is neutral, as in the kinds of cases that are

[76] D. Tuggy, 1993: 274.

used in linguistic tests: I have no attitude toward whether Tom plays checkers or not. Because of this, only one system of evaluation will activate, and crossed readings (and the like) will be hard to find. Where linguistic models must posit an error, the psychological model can make sense of the findings.

8.3 Charity

A successful diagnosis of the skeptical problem will be charitable: it should explain the skeptical problem rather than displace it, and it will make sense of individual behavior rather than attribute systematic errors. Psychological contextualism is more charitable than linguistic contextualism. Linguistic contextualism posits semantic blindness to explain the intuition of conflict. Even accepting that some people experience the intuition of conflict and some don't, as the results of studies 1 and 2 in Chapter 3 show, contextualists must still claim that some people systematically misunderstand their language. They must claim that some people tacitly fail to recognize that "know" is context-sensitive.

According to psychological contextualism the division in the intuition of conflict is not straightforwardly a mistake. We have, on my account, two systems that apply different standards of evaluation. There is an ecological rational for this, we automatically apply the permissive system to evaluating claims that do not pose a threat, and we automatically apply the stringent system to evaluating claims that do. Psychological contextualism makes no claims about what we should do, or how we should resolve a conflict between the systems when they arise. Skeptical arguments like AI induce the intuition of conflict by prompting us to apply the stringent standards and the permissive standards in evaluating a favored focal hypothesis. As I have argued, the cues that trigger these systems are well delineated. Because our cognitive powers are limited, it makes little ecological sense to subject unimportant claims, or well-supported claims, to stringent scrutiny, and so I suspect our epistemic psychology was not organized by evolution to deal with the eventuality.

Assume that there is some normative theory that tells us how this conflict should be resolved, and that that theory is invariant, implying that all people should resolve it in the same way. If correct, it would seem that psychological contextualism is no more charitable than linguistic contextualism. Where linguistic contextualism posits that some people make a mistake in understanding their language, psychological contextualism posits that some people make a mistake in reasoning.

If this is the worst challenge that can be levied at the view, then so be it. Our epistemic psychology evolved to produce ecologically rational beliefs and decisions. That we navigate the world unaffected by skeptical arguments like AI shows that our behavioral response to the intuition of conflict is ecologically rational. Thirty years of cognitive psychology shows that our reasoning departs from normative canons it myriad ways. This kind of error is not unknown. Language evolved to communicate, and to the extent that we misunderstand it, communication is impeded. Thirty years of cognitive linguistics does not show the kinds of systematic errors in understanding that linguistic contextualism must posit.

9. Bias Argument Against Skepticism

In section 3 of this chapter I also introduced one goal that a diagnosis of the skeptical problem need not achieve to be successful: it need not provide a response to the skeptic. I think that some diagnoses of the skeptical problem, like linguistic contextualism, have sacrificed empirical accuracy in favor of producing a reply to the skeptic. Linguistic contextualism is a concessive theory; it proposes that when radical skeptical hypotheses are salient we do not know quotidian perceptual facts about the world. But linguistic contextualism also safeguards empirical knowledge when skeptical hypotheses are not salient. As David Lewis sums up the position, "'I know that I have hands' [is] true in its everyday context, where the possibility of deceiving demons [is] properly ignored."[77] It is more than reasonable to think that this is a considerable theoretical virtue of linguistic contextualism. It does not do *all* we might hope for; it does not show that the skeptic's argument is fallacious, but it does save most of what we prize.

Because my project has so far been entirely descriptive, I have not offered an opinion one way or the other. Having nothing to say about whether we have empirical knowledge could be seen as a considerable disadvantage of my proposal.

Do we know we are not brains-in-vats? I have to admit that thinking about radical error possibilities so often in the course of this research has burned out my intuitions, just as linguists' acceptability intuitions are sometimes burned out by long consideration of the same linguistic construction. Sometimes I have the strong intuition that we do; skeptical

[77] D. Lewis, 1996: 564.

counter hypotheses are deeply implausible, and we have no evidence for even one instance of such deception. On other occasions I have the strong intuition that we do not; skeptical counter hypotheses are special, they are specifically chosen because we cannot rule them out.

I will not try to defend the normative conclusion that we do have everyday empirical knowledge, and in that way my proposal will compare unfavorably to linguistic contextualism. But, there are many such defenses, and I will not try to improve on the work of better philosophers. What I will argue, however, is that the empirical results presented in Chapter 4 show that radical skeptical arguments from error like AI succeed by appealing to a cognitive bias. If the competing diagnoses offered by linguistic contextualism and psychological contextualism are to be measured by weighing their normative implications, then the measurement will be of virtues of different kinds. Linguistic contextualism preserves a plausible claim by conceding a limited defeat to an implausible claim. Psychological contextualism has nothing to say about the plausible claim, but reveals a flaw in the argument for an implausible claim. I will leave it to the reader to decide which is the more significant result

9.1 The Status of Skeptical Arguments

The dialectical force of arguments like AI depends on their intuitiveness. In Chapter 1 I reviewed a large literature that shows the received view among *philosophers* is that AI is intuitive. Because they tend to find it intuitive, philosophers concede to the skeptic that the premises of AI do not need any further justification. The operating assumption of many has been that AI is intuitive because the argument appeals to nothing more than what is implicit in our ordinary epistemic practices. As I've mentioned, Michael Williams has argued that this leaves the skeptic in a powerful position: complex theoretical proposals fare poorly in comparison with intuitive arguments.[78]

We need not be so concessive to the skeptic. If the reason for conceding the high ground to the skeptic is based on the assumption they assume nothing more contentious than our ordinary epistemic practices in their arguments, then that concession is misplaced. Philosophers have relied on armchair intuitions in assessing AI. But, as we saw in Chapter 2, philosophers' intuitions are not a reliable guide to ordinary epistemic practices. I have

[78] M. Williams, 1996.

offered empirical evidence that radical versions of AI are not intuitive to most people in most contexts. This suggests that when some people in some contexts find AI intuitive, it is not a product of ordinary epistemic practice, but exceptional circumstances.

In this section I will argue that the intuitive force of radical versions of AI, when it is felt, tacitly relies on the use of an epistemic principle in a way that we would not always endorse if it were made explicit. I will call the principle the **consistency heuristic**. According to the principle, we give the same epistemic assessment to cases that are very similar. We saw that implausible counter hypotheses only generate salience effects after Cartesian priming. I propose that Cartesian priming succeeds by tacitly inducing some people to treat very implausible skeptical counter hypotheses like plausible counter hypotheses by connecting them through a series of small steps. But consistent treatment here is not an epistemic principle we should obviously accept, nor all applications of it. Indeed, I will claim that in some cases it is a cognitive bias.

Before I present the argument in more detail, a few preliminaries are in order. Even if this argument is successful, it does not amount to an argument *against* radical skepticism. There may be many arguments for radical skepticism, perhaps many good ones. I'm not offering an argument for dogmatism, only against some radical versions of AI. Moreover, I'm not offering an argument against radical versions of AI *tout court.* Cartesian priming is one way to make very implausible counter hypotheses salient, but it may not be the only way. Descartes does seem to follow this sequence in the first Meditation, introducing perceptual error first, then moving to madness and dreams before considering demon deception. There may be other ways to overcome indifference to the implausible.

Nevertheless, that most people are indifferent to implausible error possibilities does put some pressure on the skeptic to defend the first premise of AI, for it embodies a kind of claim that is not our normal epistemic practice to accept. Everything I say here leaves open the possibility that there are some methods of presentation of very implausible counter hypotheses that can produce salience effects without tacitly relying on a cognitive bias. The onus of finding one, however, lies with the skeptic.

9.2 Bias and the Appeal of AI

The plan for the rest of this section is straightforward. First I'll review how some invariantists have invoked the cognitive bias known as the availability heuristic as an

explanation of salience and as part of an argument against the skeptic. I'll offer a reason why this strategy isn't ultimately successful, irrespective of its predictions about salience effects. I'll then explain how the results of studies 2 and 3 in Chapter 4 show that the skeptic's argument only gains our assent by inducing us to tacitly deploy the consistency heuristic. I will then argue that repeated applications of this principle results in a bias in our epistemic assessments, and that this gives us good, if not overwhelming, reason to reject the form of the skeptic's argument. I'll conclude by considering the implications of the empirical study of the skeptical problem that I've undertaken here.

As we've seen already in Chapter 4, epistemic invariantists have attempted to reply to AI by arguing that salience effects in radical skeptical cases are a byproduct of the availability heuristic.[79] As Timothy Williamson puts it, "lurid stories about brains in vats, evil demons, painted mules, or gamblers who bet the farm" create "the illusion of epistemic danger" because we engender them with epistemic weight based on how easy they are to imagine, not how epistemically threatening they really are.[80]

Jennifer Nagel also offers a psychological account of salience effects, but in her case she argues they are a byproduct of epistemic egocentrism. Epistemic egocentrism is the phenomenon whereby people tend to project their beliefs, concerns, and interests onto others automatically. Nagel posits that people do not attribute knowledge in salience cases because they project their personal concern with the possibility of error onto the agent in the story, and the agent then looks hasty, distracted, or epistemically irresponsible for failing to investigate an error that has been raised. We are reluctant to attribute knowledge on this account, because the agent *seems* to violate invariant standards of good reasoning, which censure hasty, distracted, or irresponsible thoughts.[81]

Salience effects cases are used as evidence for skeptic-friendly theories of knowledge. Linguistic contextualism is one such theory we've considered at length. The anti-skeptical strategy of Nagel, Hawthorne, and Williamson is to undercut the use of salience cases by showing that their appeal depends on an error in reasoning that is the byproduct of a

[79] J. Hawthorne, 2003; T. Williamson, 2005; See also J. Vogel, 1990.

[80] T. Williamson, 2005: 226.

[81] I believe that the experiment reported in footnote 55 shows that salience effects do not depend on projection. There, I generated salience effects with likelihood judgments, and in that study there was no agent on which to project an evaluator's belief, and thus no epistemically irresponsible failure to consider the counter hypothesis. Salience effects can occur without projection.

systematic cognitive bias.[82] That we reject the principle when it is made explicit puts the burden on the skeptic to explain why we should accept it. But the strategies depend on empirical claims, if salience effects in skeptical cases are not in fact a byproduct of those biases, then there is no burden of explanation born by the skeptic. As I argued in Chapter 4, studies 2a-d show we are indifferent to implausible counter hypotheses, and so the evidence suggests we do not overestimate the epistemic danger of skeptical hypotheses. Very much to the contrary, we ignore them. The same results suggest that even if epistemic egocentrism is operative in ordinary salience cases, it is not operative in radical skeptical salience cases. Some other process seems to screen it off. We do not penalize individuals for "failing to respond to our concerns," to use Nagel's phrase, because radical skepticism does not concern us.

Experiment 3a in Chapter 4 showed that responses to salience cases exhibit order effects. In particular, this experiment showed that some people do experience salience effects for radical skeptical hypotheses when salience cases are presented in sequence from most plausible counter hypothesis to least plausible counter hypothesis. Order effects can have a number of causes. Fatigue is one: performance on the same or similar task can decline at the end of a long sequence just because experimental subjects loose interest. Practice is another: performance on the same or similar task can improve at the end of a sequence as experimental subjects learn how to perform the task.[83] I believe the order effects observed here are the result of a tacit epistemic principle I call the consistency heuristic.

According to the heuristic, when presented with similar cases we feel the intuitive need to evaluate them in the same way.[84] In experiment 3a, participants were presented with four cases, a normal case (1), a plausible error case (2), an implausible error case (3), and a very implausible error case (4). By in large, where knowledge attribution was high in (1)

---

[82] The Hawthorne-Williamson account has a surprisingly straightforward flaw. Both invoke the availability heuristic as evidence against linguistic contextualism. And while it might explain the appeal of "lurid" skeptical hypotheses in terms of a bias, it says nothing about the quotidian salience cases that contextualists typically point to. It might block radical skeptical hypotheses, but it does not undermine the contextualist's evidential base: there are no lurid stories there, only the most familiar kind of errors.

[83] In Chapter 4 I argued that the change in performance on the sequential salience task is not a byproduct of practice. Experiment 3b tested this hypothesis by presenting four very implausible salience cases in sequence, and there was no change in performance, suggesting that repeated exposure does not help participants to "recognize" something they initially missed.

[84] Discussions in the vagueness literature of the principle of tolerance resemble the principle at issue here.

(*M*=4.85), and it was very low in (2) (*M*=2.37). It is my hypothesis that some people recognized the structural similarity between (2) and (3) and felt the tacit need to evaluate them in a consistent manner. The mean response rate for (3) went up slightly (*M*=2.72), but not because the median response moved higher. Responses were categorical; participants either attributed knowledge at a high level or a low level, and not typically at an intermediate level. As in the move between (2) and (3), I believe that some individuals recognized the structural similarity between (3) and (4), and felt the tacit need to evaluate the cases in a consistent fashion. Again the mean knowledge attribution level (*M*=3.18) in (4) was higher than in (3), but this was not because individuals attributed knowledge at an intermediary level. Either they consistently denied it at the same level in (2), (3), and (4), or at some point the switched back to attributing it at a high level. Figure 8 from Chapter 4 is reproduced below to demonstrate the bimodal distribution of responses to very implausible error cases. The bi-modal distribution of results show that those participants who did experience salience effects in the radical skeptical case (4), experienced them in the same way as (2) and (3), not a moderated version of them. This consistency heuristic explains this pattern.

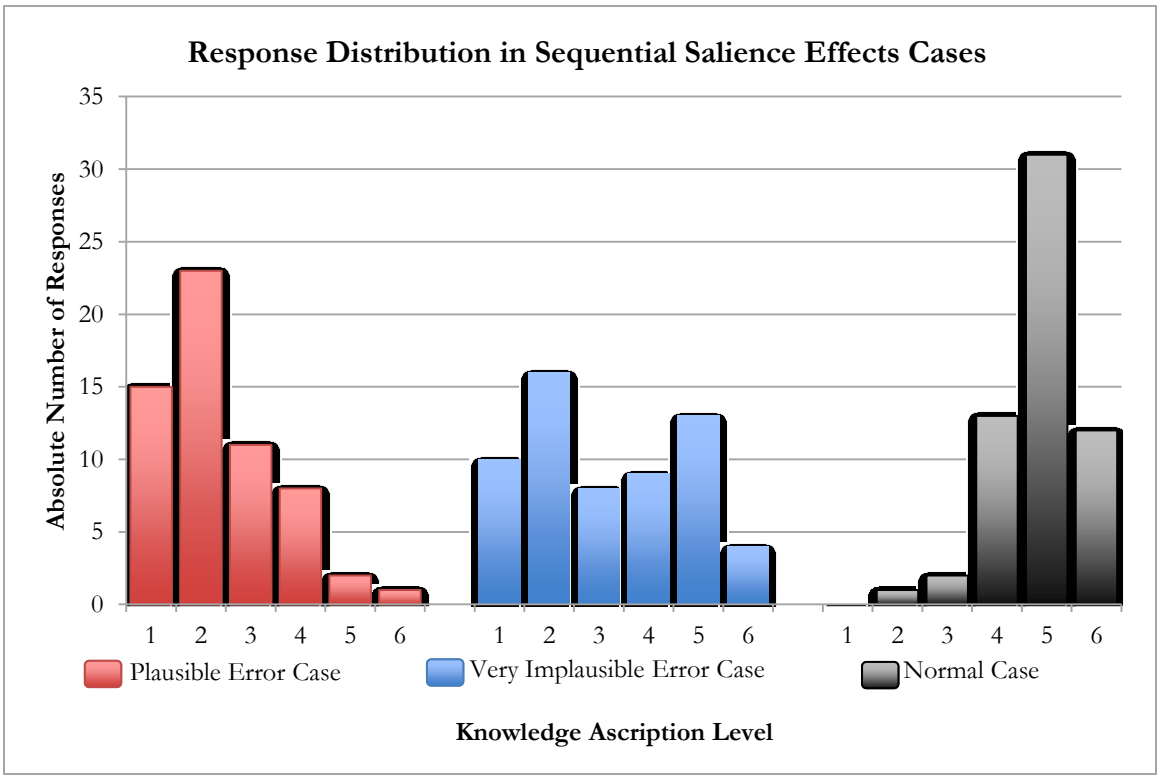


Figure 8 (chapter 4): Response distributions in sequentially presented salience effects cases illustrate the defeat structure of skeptical cases. Column counts indicate the number of participants at that level in response to the following prompt: "Please indicate the extent to which you agree or disagree with the following statement: 'John knows the bird is a Siberian grebe'." Responses were measured on

a six point likert scale where 1 = strongly disagree and 6 = strongly agree. The modal distribution in the Normal case is "5", while in the Plausible Error case it is "2". In the Very Implausible Error case, there are two modes, the first at "2" is below the midline, and the second at "5" is above the midline.

This proposal makes a prediction. Individuals who tend to prize consistency will be more prone to Cartesian persuasion. Arie Kruglanski has developed a personality scale, known as the Need for Cognitive Closure Scale (NFCS) that measures individual differences in reasoning to produce consistency. As Kruglanski describes it, individuals differ in the extent to which they "seize and freeze" on a viewpoint. High NFCS individuals are more likely to try to update their beliefs in a way that maintains consistency with prior beliefs (those they've "seized" on), and they are less likely to revise their beliefs when presented with new information (they tend "freeze" on one position).[85] A prediction of the consistency heuristic account of Cartesian priming is that individuals that are high on the NFCS scale will be more prone to Cartesian priming than those that are low on the NFCS scale. I set out to test this prediction, and the results support my claim that Cartesian priming is the result of the consistency heuristic.

Experiment 1.

Methods:

To test whether high NFCS individuals are more prone to Cartesian priming, I recruited 102 participants from the United States (age *M*=31.5; gender: 32% female; education: 40.3% bachelors degree or above; race: 76% white) on Amazon Mechanical Turk, and compensated them at minimum wage. All participants belonged to a single condition. They read all four cases from experiment 2a in chapter four in the following sequence: normal case, plausible error case, implausible error case, and very implausible error case. After reading each case participants were asked to rate their agreement to the following statement "John knows the bird is a Siberian grebe" on a six point likert scale where 1 = strongly disagree and 6 = strongly agree. On a page following the normal case, with the vignette no longer visible, participants answered a question designed to check their attention: three participants were eliminated for failing this question. On a further page, with the vignette no longer visible, participants answered the eighteen questions that make up Kruglanski's NFCS scale. Participants were divided into quartiles by NFCS score, and

[85] A. Kruglanski, 1980; A. Kruglanski & T. Freund, 1983.

salience effect levels were compared across quartiles. The results are presented in Table 5.3 below.

Experiment 1 Results:

Table 5.3 presents a clear picture of how individual differences in the need to preserve consistency results in Cartesian priming. Overall, there is an inverse correlation (r^2_s=-.198, p=.049) between NFCS score and knowledge attribution in the very implausible scenario. As NFCS increases, knowledge attribution decreases, indicating that high NFCS individuals are more likely to be affected by Cartesian priming. Comparing the performance of different NFCS quartiles reveals the clearest trends.[86] When high NFCS individuals are presented with a normal salience case, they respond as usual and they tend to deny that the agent has knowledge.[87] As more implausible salience cases are presented, however, they continue to deny the agent has knowledge. By contrast, among low NFCS individuals, as salience cases become more implausible, knowledge attribution tends to go back up. To investigate the relationship I performed a 2x3 mixed measure ANOVA, treating NFCS quartile as a between group factor, and using the error cases as within group factor. There was a main effect for NFCS group, F (1, 98)=7.05, p<.01, η^2=.214. There was the predicted interaction between NFCS group and response to sequential priming, F(2, 48)=2.50, p=.044, η^2=.124. The effect was driven primarily by the significant difference, t(50)=2.66, p=.011, in responses to the very implausible scenario, which were much lower in the high NFCS group (M=2.13, SD=1.37) than in the low NFCS group (M=3.45, SD=1.41).

[86] In what follows I'll refer to the top quartile as "high NFCS individuals" and the bottom quartile as "low NFCS individuals."

[87] Following the convention of Chapter 4, when results fall below the midline of the Likert scale, I will say that the participants "deny" the agent knowledge. When discussing the results comparatively, I will say that participants attribute knowledge at a higher or lower level, even when the level is low enough to count as denying knowledge.

	Low NFCS Quartile Mean Knowledge Ascription	High NFCS Quartile Mean Knowledge Ascription
Normal Case	4.73	5.19
Plausible Error Case	2.41	2.44
Implausible Error Case	2.82	2.50
Very Implausible Error Case	3.45*	2.13*

TABLE 15: Experiment 1: NFCS and Cartesian Persuasion. Knowledge ascription means indicate participants response to the following prompt: "Please indicate the extent to which you agree or disagree with the following statement: 'John knows the bird is a Siberian grebe'." Responses were measured on a six point likert scale where 1 = strongly disagree and 6 = strongly agree. Significant differences between NFCS quartile groups is marked by * at the $p<.05$ level; Significant differences between cases within an NFCS quartile are marked by column position: knowledge ascriptions in different column positions are significantly different at the $p<.05$ level.

The results show that the power of sequential presentation depends at least in part on individuals' need for cognitive closure. High NFCS individuals attributed knowledge in each of the three salience cases at statistically indistinguishable levels, but low NFCS individuals tended to attribute knowledge at higher levels as the skeptical counter hypothesis became more implausible. Knowledge attribution in the very implausible error cases among low NFCS individuals (M=3.45) was still lower than in the normal case (M=4.73), $F(27)=2.62$, $p=.008$. This suggests that while what I call the consistency heuristic is in play, it is not the sole mechanism at play in Cartesian persuasion. In more general terms this makes sense. Radical skepticism is a concern among philosophers, and philosophers recognize the structural similarity between quotidian error cases and the more implausible kinds at work in radical skepticism. If philosophers naturally prize consistency, or philosophical training teaches us to prize it, then it would not be surprising if philosophers were more susceptible to Cartesian priming. Indeed, it is likely they do not need it once they've encountered a plausible error case in their past. Unlike the experimental participants in this study, they do not need to be led to the skeptical result, but instead are able to recognize it own because they remember structurally similar past cases. Once we've been led to the skeptical well once, we remember the way, and we do not need a guide on subsequent occasions.

So, if the consistency heuristic is at play in explaining how skeptical arguments sometime get their grip on us, what does this imply for the skeptic's argument? Consistent treatment of similar cases is not obviously a cognitive defect. If we have good reason to evaluate a particular claim one way, then it is reasonable, though not necessary, to expect that the same reasons will apply equally to similar claims. Compare the consistency principle with plausible inequalities. If you assent to P, and Q has better epistemic credentials than P, ceteris paribus, then you should assent to Q. Likewise, if you don't assent to P, and Q has worse epistemic credentials than P, ceteris paribus, then you shouldn't assent to Q. Now compare these with the consistency principle. If you assent to P, and Q has very similar, but not equivalent epistemic credentials to P, then you should assent to Q as well. Locally this won't get you in much trouble. But applied repeatedly it leads to disaster.

The sorites paradox is a dramatic example of the kind of trouble this principle can lead to when applied repeatedly. The pattern is a familiar one, from Theseus' ship and color patch cases, to the paradox of the heap. If we accept that a 99,000 grains of sand is a heap, then we should treat 98,999 grains of sand the same, and call it a heap as well. By many easy stages we can be led to the conclusion that one grain of sand is a heap. I believe that Cartesian priming takes tacit advantage of the same principle that leads to the sorites paradox.

The paradox is a genuine one, though, for there are powerful considerations on both sides. We cannot endorse unrestricted use of the similarity principle because it would lead to widespread inconsistency. We cannot reject it for the same reason. Rejecting it seems to force us into drawing arbitrary lines between extremely similar cases. So, with the paradox of the heap, accepting the consistency principle forces us to accept that one grain of sand is a heap, and rejecting it forces us to draw arbitrary lines.

The principle can be used unobjectionably to illustrate inconsistencies in our reasoning too, i.e. *you accept P, and reject Q, but Q is very similar to P, so you should accept Q as well.* But the sorites paradox asks us to consider when we should limit repeated applications of the principle. How far is too far? I won't attempt to draw a fine line here. I will argue, thought, that when the principle is used to take us from a claim P, that we would typically accept, to a claim Q, that we would typically reject, then the use of the principle must be justified by other means.

The skeptic leads us into disaster by tacitly appealing to the consistency principle in setting up the intuitive groundwork for evaluating subjectively implausible counter hypotheses. They do so by tacitly appealing to a rule that we typically reject in other areas. If the move were made explicitly, as in the sorites paradox, each step would seem intuitively acceptable, but *the form of the argument* would no longer be intuitively acceptable. This leaves the skeptic in a weaker position than before. The received view is that each step in AI is intuitive, and the argument itself appears deductively *valid.* Once the priming steps in Cartesian persuasion are made explicitly, by contrast, while each step might also count as intuitive, the argument as a whole becomes much more questionable. If we are to accept this form of reasoning, we can without fault ask the skeptic to defend it. This is, I think, a positive result. It does not show that the skeptic is wrong, but it does reveal how the skeptic might sometimes lead us into confusion. Repeated applications of the consistency heuristic to sequentially structured cases is not part of our ordinary epistemic practice. By recognizing that the same strategy that we call into question elsewhere is at work in some skeptical arguments, we can see that skeptical arguments like AI are not as presumptionless as they might sometimes appear. If we reject the form of the argument elsewhere, irrespective of its content, we can also reject the form of the argument in the skeptical case, irrespective of *its* content.

10. Conclusion

It could be argued that I have only made a dubious kind of progress here, in effect, trading one paradox for another. But I do not believe that is the case. In chapter one I surveyed the literature on the skeptical problem, and found that by and large philosophers found it to be intuitive. Accepting that each step of arguments like AI is intuitive, and that the argument itself is apparently valid, puts the skeptic in a commanding position: they need not offer any defense of their argument. Indeed, as I've also remarked, the skeptical problem is not a confrontation with a real and genuine skeptic who actually presses and argument upon us, but a personal confrontation. In a personal confrontation, where we recognize that an argument leads to a conclusion we reject, responding to that argument depends on providing a reply that we ourselves find satisfying. But as I've also been wont to highlight before, and as Michael Williams observes, it is hard to put our minds at ease because theoretical diagnoses fare poorly against intuitive arguments.

But this is where the power of the empirical approach that I favor becomes apparent. It would be fruitless to try to meet the skeptic's arguments by pointing to empirical facts, for to do so would beg the question catastrophically. Empirical diagnosis aims at self-understanding. It aims to explain why we have the intuitions that we do, and to understand the factors that can change them. It can help us adopt an attitude of detachment whereby we do not evaluate the strength of an argument purely in virtue of the intuitive grip it has upon us. In Chapter 2 I argued that in many ways philosophers' epistemic intuitions depart from ordinary epistemic practice. I argued further that this should give us pause before taking either the philosopher's description of the skeptical problem or their diagnoses of it at face value. In Chapters 3 and 4 I conducted an empirical investigation of how ordinary epistemic practice responds to elements of the skeptical problem. In Chapter 3 we found that there are individual differences in the intuitive force of the intuition of conflict. In Chapter 4 we found that ordinary epistemic practice is to ignore implausible counter hypotheses, but that there are some ways to overcome this indifference. Taken together they validate the very general hypothesis of Chapter 2: philosopher's intuitions about the skeptical problem do differ from those of non-philosophers.

In this chapter I presented a diagnoses of the context-sensitive appeal of skeptical arguments like AI. They are not naturally intuitive to all people in all contexts, even when radical counter hypotheses have been made salient. They are, however, naturally accessible. The empirical diagnosis of the appeal of skeptical arguments can help us detach the descriptive project of understanding skeptical arguments from the normative project of explaining the conditions for knowledge. An analogy with reasoning about probabilities is instructive here. Many people find the gambler's fallacy intuitive: having met a run of bad luck they feel due a run of good luck. But the canons of probability theory dictate, in games of chance where trials are independent, that past performance does not determine anything about future performance. The intuitive appeal of the gambler's fallacy is rooted in psychological facts: facts about how we evolved and facts about how our current cognitive constitution. Normative probability theory is not true because of facts about our psychology and neither is it beholden to it. Probability theory emerged before cognitive psychology, but cognitive psychology can help us understand why our intuitions are not in accord with it.

So too, I believe, with the skeptical arguments like AI. Investigating arguments like AI empirically can help us understand their features, the conditions in which we find them

intuitive, and the cognitive mechanisms that ultimately give rise to their appeal. It is a method of investigation that merely reveals what principles are at work, answering the question of whether we should accept them is an entirely different task. Experimental philosophy doesn't undermine traditional philosophy, it is just another tool that philosopher's can use. Likewise, showing, as this investigation does, that philosopher's intuitions about the skeptical problem differ from ordinary individuals does not show that one group is right, or that one group is wrong. What it does do is level the playing field. It reveals the principles of our epistemic psychology and it leaves open, to arguments on both sides, whether they are correct.

Bibliography

Alexander, J. (2012). *Experimental Philosophy: An Introduction*. Cambridge, MA: Polity Press.

Alexander, J., C. Gonnerman, & J. Waterman (forthcoming) "Salience and Epistemic Egocentrism," in J. Beebe (Ed.) *Advances in Experimental Epistemology,* New York: Bloomsbury Academic.

Arkes, H. (1991). "Costs and benefits of judgment errors: Implications for debiasing," *Psychological Bulletin*, 110: 486–498.

Austin, J. L. (1946). "Other minds." *Proceedings of the Aristotelian Society*, 20: 149–187.

Baron, J. (1995). "Myside bias in thinking about abortion," *Thinking and Reasoning, 1*, 221-235.

Baron, J. (2008). *Thinking and Deciding*. New York: Cambridge University Press.

Bealer G. (1992). "The Incoherence of Empiricism," *Proceedings of the Aristotelian Society*, Supplementary Volumes, Vol. 66: 99-143.

Beebe, J. and W. Buckwalter (*forthcoming*). "The Epistemic Side-Effect Effect," *Mind & Language.*

Behrend, Tara, David Sharek, Adam Meade, & Erice Wiebe (2011). "The Viability of Crowdsourcing for Survey Research," *Behavioral Research* 43: 800 – 813.

Bengson, J. (2010). *The Intellectual Given,* Unpublished Ph.D. dissertation submitted to the Faculty of Arts and Sciences at the University of Texas-Austin.

Bengson, J. (2013) "Experimental Attacks on Intuitions and Answers," *Philosophy and Phenomenological Research.* Vol. 86, No. 3: 495-532.

Boghossian, P. (2000). "Knowledge of Logic," in *New Essays on the A Priori*, P. Boghossian and Christopher Peacocke (Eds.), New York: Oxford University Press.

Bostrom, N. (2003). "Are you living in a computer simulation." *Philosophical Quarterly*, Vol. 53, No. 211: 243-255

Bourget, D. and Chalmers, D. (forthcoming) "What Philosopher's Believe," *Philosophical Studies*.

Buckwalter, W. (2010). "Knowledge isn't closed on Saturday," *Review of Philosophy and Psychology*, 1: 395-406.

Buckwalter, W. (2014). "The mystery of stakes and error in ascriber intuitions', in J. Beebe (Ed.), *Advances in Experimental Epistemology*, London: Continuum Press.

Buckwalter, W., & Stich, S. (forthcoming). "Gender and philosophical intuition." In J. Knobe & S. Nichols (Eds.) Experimental philosophy (Vol. 2). Oxford: Oxford University Press.

Cappelen, H. (2012) *Philosophy Without Intuitions*. New York: Oxford University Press.

Cappelen, H. & E. Lepore (2005). *Insensitive Semantics: A Defense of Semantic Minimalism and Speech Act Pluralism.* Oxford: Blackwell

Chaiken, S. & Y. Trope (1999). *Dual Process Theories in Social Psychology.* New York: Guildford Books.

Cheng, P. W., Holyoak, K. J., Nisbett, R. E., & Oliver, L. M. (1986). "Pragmatic versus syntactic approaches to training deductive reasoning." *Cognitive Psychology*, 18, 293-328.

Cherniak, C. (1991). *Minimal Rationality.* New York: Bradford Books.

Cohen, S. (1988). "How to be a fallibilist," *Philosophical Perspectives*, 2: 91–123.

Cohen, S. (1999). "Contextualism, skepticism, and the structure of reasons." *Philosophical Perspectives,* 13: 57–89.

Cohen, S. (2001). "Contextualism Defended," *Philosophical Studies*, 103: 87-98.

Cohen, S. (2002). "Basic Knowledge and the Problem of Easy Knowledge," *Philosophy and Phenomenological Research,* 65: 302 – 329.

Cohen, S. (2004). "Contextualism and Unhappy-Face Solutions: Reply to Schiffer," *Philosophical Studies,* 119: 185-97.

Cullen, S. (2010). "Survey Driven Romanticism," *Review of Philosophy and Psychology*. Vol. 1 (2): 275-296

Dawson, E., Gilovich, T. & Regan, D.T. (2002). "Motivated Reasoning and Performance on the was on Selection Task," *Personality and Social Psychology Bulletin* 28, 1379-1387

DeRose, K. (1992). "Contextualism and Knowledge Attributions," *Philosophy and Phenomenological Research*, vol 52: 913 – 29.

DeRose, K. (1995). "Solving the Skeptical Problem," *Philosophical Review*, 104: 1-52.

DeRose, K. (2009). *The Case for Contextualism.* New York: Oxford University Press.

DeRose, K. (2011). "Contextualism, Contrastivism, and X-Phi Surveys," *Philosophical Studies* 156: 81 – 110.

Descartes, R. (2006). *Meditations on First Philosophy in Which the Existence of God and the Distinction between the Soul and the Body are Demonstrated, in Meditations, Objections, and Replies*. Roger Ariew and Donald Cress Translators and Eds., New York: Hackett.

Devitt, M. (1991). *Realism and Truth* (2nd Ed.), Princeton, NJ: Princeton University Press.
Ditto, P., & D. Lopez (1992). "Motivated Skepticism: Use of Differential Decision Criteria for Preferred and Nonpreferred Conclusions," *Journal of Personality and Social Psychology*. Vol. 63, No. 4: 568-584.

Dougherty, T. & P. Rysiew (2009). "Fallibilism, Epistemic Possibility, and Concessive Knowledge Attributions," *Philosophy and Phenomenological Research.* Vol. 78 (1):123-132.

Dretske, F. (1970). "Epistemic Operators," *Journal of Philosophy*, vol. 67: 1007 – 1023.

Dretske, F. (1981). "The Pragmatic Dimension of Knowledge," *Philosophical Studies*, vol. 40: 363 – 378.

Dunbar, R. (2004). *The Human Story.* London: Faber & Faber.

Edwards, K. & E. Smith (1996) "A Disconfirmation Bias in the Evaluation of Arguments," *Journal of Personality and Social Psychology*. Vol. 71, No. 1.5-24

Feltz, A. and Zarpentine, C. (2010). "Do you know more when it matters less?' *Philosophical Psychology*, 23: 683 - 706.

Field, A. (2009). *Discovering Statistics Using SPSS*, 3rd Edition. Sage Publications: London.

Fischhoff, B. (1982). "Debiasing," In D. Kahneman, P. Slovic, & A. Tver- ky (Eds.), *Judgment under uncertainty: Heuristics and biases:* 422-444. New York: Cambridge University Press.

Fiske, S. T, & S. L. Neuberg (1990). "A continuum model of impression formation from category-based to individuating processes: Influences of information and motivation on attention and interpretation." In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1-74). San Diego, CA: Academic Press.

Foot, P. (1967). "The problem of Abortion and the Doctrine of the Double Effect," *Oxford Review*, no. 5.

Frankland, B. W. & B. Zumbo (2002). "Quantifying Bimodality Part I: An Easily Implemented Method Using *SPSS*," *Journal of Modern Applied Stastical Methods*, Vol. 1, No. 1: 157 – 166.

Frederick, S. (2005). "Cognitive reflection and decision making," *Journal of Economic Perspectives*, 19(4), 25-42

Gettier, E. (1963) "Is Justified True Belief Knowledge?", *Analysis*, vol. 23, No. 6: 121 - 123.

Gigerenzer, G. (2002). *Adaptive Thinking: Rationality in the Real World.* New York: Oxford University Press.

Gigerenzer, G.. & R. Selten (2002). "Rethinking Rationality," in G. Gigerenzer and R.Selten, (Eds.) *Bounded Rationality: The Adaptive Toolbox.* Cambridge, MA: MIT Press.
Gilovich,
Goldman, A. (2007). "Philosophical intuitions: Their target, their source, and their epistemic status," *Grazer Philosophische Studien* Vol. 74: 1 – 26.

Goldman, A. (2010). "Philosophical naturalism and intuitional methodology," (Romanell Lecture), *Proceedings and Addresses of the American Philosophical Association*:115 – 150.

Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., & Cohen, J.D. (2001). "An fMRI investigation of emotional engagement in moral Judgment," *Science*, Vol. 293, 2105-2108.

Grice, H.P. (1957). "Meaning," *Philosophical Review*, 66: 377–388.

Gross, S. (2001). *Essays on Linguistic Context-Sensitivity and its Philosophical Significance*, New York: Routledge.

Haidt, J., S. Koller & M. Dias (1993). "Affect, Culture and Morality," *Journal of Personality & Social Psychology*, 65, 4: 613 - 628.

Haidt, Jonathan 2001. "The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment," *The Psychological Review* Vol. 108: 814 - 834.

Hales, S. (2006) *Relativism and the Foundations of Philosophy*. Cambridge, MA: MIT Press.

Hansen, N. & Emmanuel C. (forthcoming). "Experimenting on Contextualism," *Mind and Language.*

Harman, G. (1968). "Knowledge, inference, and explanation," *American Philosophical Quarterly* Vol. 5: 164 – 173.

Hauser, M., Cushman, F., Young, L., Kang-Xing Jin, R. And Mikhail, J. (2007). "A Dissociation Between Moral Judgments and Justifications*," Mind & Language*, 22: 1–21

Hawthorne, J. (2004). *Knowledge and Lotteries*. New York: Oxford University Press.

Hazlett, A. (2014). *A Critical Introduction to Skepticism*. New York: Bloomsbury Academic.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Gintis, H., McElreath, R., and E. Fehr (2001). "In search of Homo economicus: Experiments in 15 Small-Scale Societies," *American Economic Review*, 91(2): 73-79.

Henrich, J., Heine, S. & A. Norenzayan (2010). "The Weirdest People in the World?," *Behavioral and Brain Sciences*. 33(2-3): 61-83; discussion 83-135.

Hintikka, J. (1999) "The Emperor's New Intuitions," *The Journal of Philosophy*, Vol. 96, No. 3: 127 – 147.

Hirt, E. & K. Markman (1995). "Multiple Explanation: A Consider-an-Alternative Strategy for Debiasing Judgments," *Journal of Personality and Social Psychology*. Vol. 69, No. 6: 1069-1086.

Horowitz, I. A. (1997). "Reasonable doubt instructions: Commonsense justice and standard of proof." *Psychology, Public Policy, and Law*, 3, 285–302.

Hume, David (1978). *A Treatise of Human Nature*, New York: Oxford University Press.
Ichikawa, J. 2011. "Quantifiers and Epistemic Contextualism," *Philosophical Studies* 155:383 – 398.

Johnson-Laird, P. N. & Wason, P. C. (1970) "Insight into a logical relation." *Quarterly Journal of Experimental Psychology. Vol. 22(1): 49–61.*

Jost, J. T., Hennes, E. P., & Lavine, H. (in press). "'Hot' political cognition: Its self-, group-, and system- serving purposes." In D. E. Carlson (Ed.), *Oxford handbook of social cognition*. New York: Oxford University Press.

Kahan, D., E. Peters, E. Dawson, & P. Slovic (forthcoming). "Motivated Numeracy and Enlightened Self-Government," Yale Culture and Cognition Project, Working Paper No. 116.

Kahneman, D. (2011) "Thinking Fast and Slow," New York: Farrar, Straus and Giroux

Kahneman, D., & A. Tversky (1973). "On the psychology of prediction," *Psychological Review* 80: 237–251.

Kahneman, D. & A. Tversky (1983). "Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment," *The Psychological Review*. Vol. 90 No. 4.

Kahneman, D., & A. Tversky (1985). "Evidential impact of base rates." In D. Kahneman, P. Slovic & A. Tversky (Eds.). *Judgment under uncertainty: Heuristics and biases*: 153–160.

Kassarjian, H. H., & Cohen, J. B. (1965). "Cognitive dissonance and consumer behavior," *California Management Review,* 8*,* 55-64.

Klayman, I, & Ha, Y.-W (1989). "Hypothesis testing in rule discovery: Strategy, structure, and content," *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 15*,* 596-604.

Knobe, J. & J. Schaffer (2012). "Contrastivism Surveyed," *Nous.*Vol. 46 (4): 675-708.

Knobe, J. & S. Nichols (2008) "An Experimental Philosophy Manifesto," in Experimental Philosophy, Joshua Knobe & Shaun Nichols (eds.), New York: Oxford University Press.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). "Reasons for confidence," *Journal of Experimental Psychology: Human Learning and Memory,* 6,107-118.

Kripke, S. (1991). Naming and Necessity. New York: Wiley Blackwell.

Kruglanski, A. W (1980). "Lay epistemology process and contents," *Psychological Review,* Vol. 87: 70-87.

Kruglanski, A. W, & T. Freund (1983). "The freezing and unfreezing of lay-inferences: Effects on impressional primacy, ethnic stereotyping, and numerical anchoring," *Journal of Experimental Social Psy- chology,* Vol. 19: 448-468.

Kunda, Z. (1987). "Motivated Inference: Self-serving Generation and Evaluation of Causal Theories." Journalof Personality and Social Psychology Vol. 53(4): 636-47.

Kunda, Z. (1990). "The Case for Motivated Reasoning," *Psychological Bulletin* Vol. 108, No. 3: 480-498.

Larrick, R.P. (2004). "Debiasing." In D.J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making.* Oxford, United Kingdom: Blackwell: 316– 338.

Lehrer, K. (1990). *The Theory of Knowledge.* (2nd Ed.), Boulder, CO: Westview Press.

Lepore, Ernie & Herman Cappelen 2003. "Context Shifting Arguments", *Philosophical Perspectives* 17 (1): 25-50.

Lewandowsky, S., U. Ecker, C. Seifert, N. Schwarz, & J. Cook (forthcoming). "Misinformation and its Correction: Continued Influence and Successful Debiasing." *Psychological Sciences in the Public Interest.*

Lewis, D. (1979). "Scorekeeping in a language game," *Journal of Philosophical Logic*, 8: 339 – 359.

Lewis, D. (1996). "Elusive Knowledge," *Australasian Journal of Philosophy*, Vol. 74, No. 4.

Lord, C. G., M. Lepper, & E. Preston. (1984). "Considering the opposite: A corrective strategy for social judgment." *Journal ofPersonality and Social Psychology,* Vol. 47: 1231-1243.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). "Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence," *Journal of Personality and Social Psychology,* Vol. 37: 2098-2109.

Ludwig, K. (2007). "The Epistemology of Thought Experiments: First vs. Third Person Approaches," in *Midwest Studies in Philosophy, Philosophy and the Empirical*, 31 (1): 128–159.

Marcus, G. (2009). *Kludge: The Haphazard Evolution of the Human Mind.* New York: Norton-Psychology.

May, J., W. Sinnott-Armstrong, J. Hull, & A. Zimmerman (2010). "Practical interests, relevant alternatives, and knowledge attributions: an empirical study," *Review of Philosophy and Psychology*, 1: 265 - 273.

McCloskey, M. (1983). "Intuitive Physics," Scientific American, Vol. 248, No. 4: 122 - 130.

Mercier, H. & D. Sperber (2011). "Why Do Humans Reason? Arguments for and Argumentative theory," *Behavioral and Brain Sciences.* Vol. 34: 57-111.

Miller, D. T, & Ross, M. (1975). "Self-serving biases in attribution of causality: Fact or fiction?" *Psychological Bulletin, 82,* 213-225.

Montminey, M. (2009). "Contextualism, Invariantism and Semantic Blindness", *Australasian Journal of Philosophy*, 84, 4: 639 – 657.

Murphy, L. M. (2010). *Lexical Meaning.* New York: Cambridge University Press.

Nagel, J. (2007) "Epistemic Intuitions," *Philosophy Compass.* 2/6: 792 - 819.

Nagel, J. (2012a) "Intuitions and Experiments: A defense of the case method in epistemology," *Philosophy and Phenomenological Research*, Vol. 85 No. 3: 495 – 527.

Nagel, J. (2012b) "Knowledge ascriptions and the psychological consequences of thinking about error," *Philosophical Quarterly*, 60(239), 286–306.

Nichols, S. & A. Pinillos (forthcoming) "Infallibilism and Bayesian Updating."

Nisbett, R., & Y. Miyamoto (2005). "The influence of culture: holistic versus analytic perception," *Trends in Cognitive Sciences*, Vol. 9, No. 10: 467 – 473.

Nisbett, R., K. Peng, I. Choi, & A. Norenzayan (2001). "Culture and systems of thought: Holistic versus analytic cognition," *Psychological Review*, 108: 291-310

Norenzayan, A., & R. Nisbett (2000). "Culture and causal cognition," *Current Directions in Psychological Science*, 9: 132-135.

Norenzayan, A., E. Smith, & B. Kim, & R. Nisbett (2002). "Cultural preferences for formal versus intuitive reasoning," *Cognitive Science*, 26: 653-684.

Norenzayan, A., R. Nisbett, E. Smith, & B. Kim (1999). "Rules vs. similarity as a basis for reasoning and judgment in East and West," Unpublished Manuscript. Ann Arbor: University of Michigan.

Oliver, J. E., & T. Wood (2012). "Conspiracy theories, magical thinking, and the paranoid style(s) of mass opinion," University of Chicago Working Papers Series. URL: http://political-science.uchicago.edu/faculty-workingpapers/ (retrieved, June 2014).

Oswald, M. E., & Grosjean, S. (2004). "Confirmation bias," In R. F. Pohl (Ed.). *Cognitive Illusions. A Handbook on Fallacies and Biases in Thinking, Judgement and Memory*. Hove and N.Y.: Psychology Press

Papineau, D. (2011) "What is x-phi good for?" *The Philosophers' Magazine*, 52: 83-88.

Partee, B. (2004). "Comments on Jason Stanley's 'On the Linguistic Basis for Contextualism'," *Philosophical Studies,* 119: 147 – 159.

Petty, R., Cacioppo, J. and Goldman, R. (1981). "Personal involvement as a determinant of argument-based persuasion," *Journal of Personality and Social Psychology*, 41(5): 847–855.

Pinillos, N. Á. (forthcoming). "Knowledge, Experiments and Practical Interests," *New Essays On Knowledge Ascriptions,* Jessica Brown and Mikkel Gerken (Eds.) New York: Oxford University Press.

Pohl, R. & W. Hell (1996). "No Reduction in Hindsight Bias after Completing Information and Repeated Retesting," *Organizational Behavior and Human Decision Processes*. Vol. 67: 49-58

Pritchard, D. 2002. "Recent Work on Radical Skepticism," *American Philosophical Quarterly,* Vol. 39, no. 3: 215 - 257.

Recanati, F., 2004. *Literal meaning*, Cambridge: Cambridge University Press.

Reed, B. (2012). "Fallibilism," *Philosophy Compass*. 7/9 (2012): 585-96.

Ross, L., Lepper, M. R. & Hubbard, M. (1975) "Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm," *Journal of Personality and Social Psychology*. Vol. 32(5): 880–92.

Ross, M., McFarland, C. & Fletcher, G. J. (1981) "The effect of attitude on the recall of personal histories," *Journal of Personality and Social Psychology.* Vol. 40(4): 627– 34.

Rozeboom, W. (1967) "Why I Know So Much More than You Do," *American Philosophical Quarterly*, Vol. 4, No. 4: pgs. 281 – 290.

Schaffer, J. (2005a). "Contrastive knowledge." *Oxford Studies in Epistemology*, 1, 235–271.

Schaffer, J. (2005b). "What shifts? Thresholds, standards, or alternatives?" In G. Preyer & G. Peter (Eds.), *Contextualism in Philosophy* (pp. 115–130). Oxford: Oxford University Press.

Schaffer, J. 2004. "From Contextualism to Contrastivism," *Philosophical Studies*. Vol. 119: 73–103.

Schiffer, S. (1996). "Contextualist Solutions to Scepticism", *Proceedings of the Aristotelian Society* 96: 317-33.

Schwarz, N. (1995). "What respondents learn from questionnaires: The survey interview and the logic of conversation," *International Statistical Review/Revue Internationale de Statistique,* 63(2):153–168.

Schwarz, Norbert. 1996. *Cognition and communication: Judgmental biases, research methods, and the logic of conversation*. Mahwah: Erlbaum.

Schwitzgebel, E. & F. Cushman (forthcoming) "Professional Philosophers' Susceptibility to Order Effects and Framing Effects in Evaluating Moral Dilemmas."

Shope, R. (1983). *The Analysis of Knowing: A Decade of Research*, Princeton, NJ: Princeton University Press.

Shweder, R. A. (1977). "Likeness and likelihood in everyday thought: Magical thinking in judgments about personality." *Current Anthropology* **18** (4): 637–58 (637).

Sloman, S. A. (1996). "The empirical case for two systems of reasoning," *Psychological Bulletin* Vol. 119(1): 3–22.

Sperber, D. (2001). "An evolutionary perspective on testimony and argumentation," *Philosophical Topics*. Vol. 29: 401–13.

Stanley, Jason 2004. "On the linguistic Basis for Contextualism," *Philosophical Studies*, Vol. 119, No. 1/2, Contextualism in Epistemology and Beyond: 119-146.

Stanovich, K. E. & West, R. F. (2007). "Natural myside bias is independent of cognitive ability." *Thinking & Reasoning* 13(3): 225–47.

Stanovich, K. E. & West, R. F. (2008). "On the failure of cognitive ability to predict myside and one-sided thinking biases." *Thinking & Reasoning* 14(2): 129–67.

Stanovich, K. R, F. West, & M. E. Toplak (2013). "Myside Bias, Rational Thinking, and Intelligence," *Current Directions in Psychological Science*. Vol. 22(4): 259–264.

Starmans, C. & Friedman, O (2012). "The Folk Conception of Knowledge," *Cognition*, Wl. 124: 272 - 283.

Stein, E. (1996) *Without Good Reason.* New York: Oxford University Press.

Stich, S. (1993). *The Fragmentation of Reason*. New York: Bradford Books

Stich, S. (2001) "Plato's Method Meets Cognitive Science," *Free Inquiry*, 21, 2, Spring: 36-38.

Stich, S. (2006). "Is Morality an Elegant Machine or a Kludge?" *Journal of Cognition and Culture* 6 (1-2):181-189.

Stich, S. (2012) "Do Different Groups Have Different Epistemic Intuitions? A Reply to Jennifer Nagel," *Philosophy and Phenomenological Research*.

Stine, G. (1976). "Skepticism, Relevant Alternatives, and Deductive Closure." *Philosophical Studies* 29: 249-261.

Stroud, Barry 1984. *The Significance of Philosophical Skepticism*. New York: Oxford University Press.

Swain, S., J. Alexander, & J. Weinberg (2008). "The instability of Philosophical Intuitions - Running Hot and Cold on Truetemp," *Philosophy and Phenomenological Research*. Vol. 77 No. 1: 138 - 155.

Tenney, E., H. Cleary, & B. Spellman (2009a). "Unpacking the doubt in 'Beyond a reasonable doubt:' Plausible alternative stories increase not guilty verdicts." *Basic and Applied Social Psychology, 31,* 1-8.

Tenney, E., H. Cleary, & B. Spellman (2009b). "This Other Dude Did it!" *The Jury Expert: The Art and Science of Litigation Advocacy"* July: 37-56.

Tooby, J. & L. Cosmides (2005). "Resolving the debate on innate ideas: Learnability constraints and the evolved interpenetration of motivational and conceptual functions." In P. Carruthers, S. Laurence, S. & Stich (Eds.), *The Innate Mind: Structure and Content*. NY: Oxford University Press.

Toplak, M., West, R. and Stanovich, K. (2011). "The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks," *Memory & Cognition* 39(7), 1275-1289.

Travis, C. 1997. "Pragmatics," In B. Hale and C. Wright (Eds.), *A Companion to the Philosophy of Language*: 87 - 106. Oxford: Blackwell.

Tuggy, D. (1993). "Ambiguity, Polysemy, and Vagueness," *Cognitive Linguistics* 4-3: 273-290

Tversky, A., and D. Kahneman, 1973. "Availability: a Heuristic for Judging Frequency and Probability", *Cognitive Psychology*, 5: 207 - 32.

Unger, P. (1971). "A defense of skepticism," *The Philosophical Review*, Vol. 80(2): pgs. 198 – 219.

Unger, P. 2002 (Orig. ed. 1975). *Ignorance: A Case for Skepticism*. New York: Oxford University Press.

van Woudenberg, R. 2005. "Contextualism and the Many Senses of Knowledge," *Grazer Philosophische Studien*, 69: 147-164.

Vogel, J. (1990). "Cartesian Skepticism and inference to the best explanation," *Journal of Philosophy*. Vol. 87(11): 658-666.
Wason, P. C. (1960). "On the failure to eliminate hypotheses in a conceptual task." *Quarterly Journal of Experimental Psychology,* Vol. 12*,* 129-140.

Weinberg, J. (2007). "How to Challenge Intuitions Empirically Without Risking Skepticism," *Midwest Studies In Philosophy*, 31: 318–343.

Weinberg, J., S. Nichols, & S. Stich (2001) "Normativity and Epistemic Intuitions," *Philosophical Topics*, 29, 1 & 2: 429-460.

Weiner, M. (2009). "The Mostly Harmless Inconsistency of Knowledge Ascriptions," *Philosophers' Imprint,* Vol.9, No. 1: 1-25.

Wells, G. L., Rydell, S. M., & Seelau, E. P. 1993. "The selection of distractors for eyewitness lineups," *Journal of Applied Psychology,* 78*:* 835– 844.

Williams, Bernard 1978. *Descartes: The Project of Pure Enquiry.* New Jersey: Pelican Books, Humanities Press.

Williams, Michael 1995. *Unnatural Doubts.* Princeton University Press: Princeton, NJ.

Williamson, T. (2005a) "Armchair philosophy, metaphysical modality and counterfactual thinking," *Proceedings of the Aristotelian Society*, 105: 1 – 23.

Williamson, T. (2005b). "Contextualism, subject-sensitive invariantism, and knowledge of knowledge," *The Philosophical Quarterly*, vol. 55: 213 - 235.

Williamson, T. (2007) The Philosophy of Philosophy. New York: Wiley Blackwell.

Windschitl, P. D. & J. R. Chambers (2004). "The Dud-Alternative Effect in Likelihood Judgment," *Journal of Experimental Psychology*: *Learning, Memory, and Cognition.* Vol. 30, No. 1: 198-215.

Curriculum Vita

John Philip Waterman was born in Peterborough New Hampshire on March 6, 1978. In 2000 he received a B.A. in philosophy from Vassar College. In 2006 he received a M.Sc. in the history and philosophy of science from the London School of Economics and Political Science. In the same year he began his Ph.D. studies at The Johns Hopkins University. He has taught philosophy at Johns Hopkins, Stevenson University, and Colby College.